PSYCHOLOGY IN RUSSIA: STATE OF THE ART

Volume 10 Issue 3 2017

Special issue CONTEMPORARY COGNITIVE SCIENCE

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society | Lomonosov Moscow State University

Special issue CONTEMPORARY COGNITIVE SCIENCE

Cognitive science: The art and its implications

Boris M. Velichkovsky

National Research Center "Kurchatov Institute", Moscow, Russia
M.V. Lomonosov Moscow State University, Moscow, Russia
Russian State University for the Humanities, Moscow, Russia
Moscow Institute for Physics and Technology, Moscow, Russia
Technische Universität Dresden, Germany

Contemporary cognitive science is the latest version of the century-long quest for a better understanding of the human mind and brain. Various disciplines have brought together empirical methods and theoretical models from their fields of study to further this effort. This multidisciplinary convergence widely known today by the acronym NBIC(S)[1] is the general context for the present endeavor.

Why is a journal with "Psychology" in its title concerned with this development? In fact, those were the founding fathers of psychology, physicist Hermann von Helmholtz and physician Wilhelm Wundt, who demonstrated the conviction that psychology as science can only be one of interdisciplinary kind. In the 20th Century, Jean Piaget gave psychology a central place among the sciences, because, in his view, only psychology studies the conscious mind that makes science and critical thinking possible. But he insisted that he was not a psychologist, explaining that he studied "epistemic" rather than "psychological" issues. In a similar vein, some psychologists would say today that they rather belong to a multidisciplinary cognitive community whereby adjective "cognitive" replaces here what Piaget called "epistemic" half a century ago.

This is the case of authors participating with their papers in this special issue of *Psychology in Russia*. Many of them are not psychologists by training, and they do their professional work in fields as varied as mathematical physics, neuroimaging, molecular biology, and the pragmatics of communication. Nevertheless, their con-

[1] As in the name of the Complex of NBICS-technologies at the National Research Center "Kurchatov Institute", where the capitalized letters stay for Nanotechnology, Biotechnology, Information technology, Cognitive, and Social sciences.

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0300
http://psychologyinrussia.com

tributions to the journal's topic, and, in my opinion, to the future of psychology as a science, are preeminent. Overall, in order to solve our century-long problems, we have to look to a diversity of approaches, and combine them in a convergent way. This is the first implication to be drawn from contemporary cognitive science.

The second implication is that emphasis must be placed on application and technologies. It is no accident that NBICS studies deal with technologies, not endless verbal or even experimental exercises in an ivory tower. Cognitive psychology, with its bias for academic sterility, was for too long deprived of almost any practical significance. But the proof of the "cognitive pudding" is in the practical "eating." The emphasis on applications does not mean the degradation of basic research. There are numerous links between basic and applied cognitive studies, which once again show that there is nothing more practical than a good theory. It can easily be demonstrated, from the research being done on perception, attention, memory, and communication, that every scientifically established fact about the organization of human cognition has important practical consequences. If not, then we simply were yet unable to recognize how the results could be applied in the major domains of praxis — engineering, medicine, and education.

The environments in which we carry out our everyday activities are complex, dynamic, sometimes dangerous, and demanding demanding a lot of social of social skills — even wisdom. From a purely quantitative point of view, they are producing streams of data that vastly exceed the known limits of human cognitive capacities. However, in striking contrast to existing "intelligent" technical systems, we, as a rule, easily understand and intelligently act in everyday situations. How do people use their perplexed brain machinery and notoriously limited cognitive resources to manage ever increasing environmental demands? The topics dealt with in this special issue have a more or less direct relationship to the answer to this question. We suspect that the general answer is called "consciousness," with its variety of derived abilities such as understanding, intellect (at least, the fluent version of it), and higher-order emotions and feelings.

Heinz Heckhausen once remarked that scientific psychology first banished the soul from its lexicon, and then lost consciousness and the ability to reason. With the rise of cognitive psychology in the second half of the 20th Century, reasoning and other cognitive functions have been returned from their behaviorist exile. Today, we see an unprecedented expansion of cognitive and neurocognitive studies of consciousness.[1] Hope is stronger than ever that this is the time to tackle big research questions. Perhaps, this is an illusion, like consciousness itself, but the emphasis on consciousness and volition is clearly seen in the collection of articles in this issue. Among the articles, there are research papers directly dedicated to phenomenology and to the brain's mechanisms of consciousness and, symptomatically, to practical applications in such domains of new and emerging technologies as virtual reality ("feeling of presence") and eye-brain-computer interfaces ("intentions," "sense of agency").

The appeal to ask questions that could not be answered for decades and centuries is of limited value unless we have powerful research tools for reaching the

1 By way of analogy, one can ask about the possible return of the soul in scientific discourse. This is perfectly conceivable but not before both trends — cognitive and affective research — are united in something like cognitive-affective science (see, e.g., Panksepp, 1998).

answers. Here we are again at a crossroads, with NBICS-convergence being the basis for the development of such research methods: from non-invasive brain imaging and eye-tracking instruments with their specialized mathematical tools, to the methods of genomic, epigenetic and metabolomic studies revealing underlying molecular mechanisms.

This convergence of methodologies seems to lead to a convergence in results. To be more specific, I will address several implications of the recent discovery which demonstrated a profound asymmetry in effective (cause-and-effect) connections between the left and right hippocampi within the human default mode network, or DMN (Ushakov et al., 2016). The latter is a set of functionally and structurally connected brain areas that are activated at rest and deactivated by external stimulation. The lateralization of the neocortex has been a well-established fact since the famous observation by Paul Broca that the left hemisphere plays a crucial role in speech production ("*Nous parlons avec l'hémisphère gauche*"). However, our finding is a *novum* because, for the first time, it refers to an asymmetry in the relationship between subdivisions of the limbic system, i.e. relatively deep structures of the paleocortex.

Moreover, the pattern of causal relationships, which is characteristic for the right hippocampus, seems to be of particular importance for self-referential cognition and higher-order forms of emotional life. The key feature of this pattern is a holistic representation of one's egocentric surroundings. With the centrality of Ego/Self in such a representation, it can provide an easy-to-find gateway into much of what we used to call "subjective experience." The access to the gateway seems to be open both for prefrontal cortices and for tertiary associative structures around right temporoparietal junction in the posterior part of the human brain (Velichkovsky et al., 2017).

Several works in this collection of papers seem to support this hypothesis in a number of research domains, at least indirectly. All of them highlight the rightward lateralization of brain mechanisms for self-referential processing. Besides the phenomenology of everyday problem-solving and neuropsychological testing in patients with unilateral brain lesions, this research includes neuroimaging studies of reading, remembering, and cognitive involvement in motor behavior. Last but not least, a pronounced difference in the gene expression within the left and right parts of the human frontopolar cortex is reported (Dolina et al., 2017). Here again, a strong right-sided predominance in the differential expression of protein-coding genes was established.[1] Many of these differentially expressed genes are known for their role in hippocampal formation and, if considered in a clinical context, for their relationship to schizophrenia.

Although basic mechanisms of language perception and production, localized primarily in the left hemisphere, have been traditionally considered as the *differentia specifica* of *Homo sapiens*, parts of the right hemisphere have a larger volume and demonstrate more rapid growth, both in anthropogenesis and in early ontogenesis. This fact can be related to the Yakovlevian Torque phenomenon, in which frontal structures anterior to the right Sylvian fissure are "torqued forward" relative to their counterparts on the left. This phenomenon has been supported

1 The pattern of asymmetry in up-expression is less apparent, missing, or even completely reversed toward leftward asymmetry if we are considering short non-coding molecules, i.e. microRNA and its precursors (Nedoluzhko et al., in preparation).

by fragmented paleo-neurological findings dating back nearly one million years (Hrvoj-Mihic, Bienvenu, Stefanacci, Muotri, & Semendeferi, 2013). It looks as if we are unique among species not for our language ability but rather for our elaborated mechanisms of controlling what we are saying and doing. Thus, the upper-level mechanisms may precede and definitively influence, in a top-down manner, the initial phases in evolution of human language (Deacon, 1996).

Instead of going into details on this multilevel brain-and-mind architecture (see, e.g., Velichkovsky, 1999, 2002), I will only stress the coherence among the empirical findings. The right hemisphere was for decades considered as subordinate and undeveloped, at best specialized in spatial information processing. However, the new findings show that it is crucially important for self-referential cognition, theory of mind, and higher forms of emotional comprehension. An unexpected result also is that not only prefrontal mechanisms are involved, but posterior structures near the temporoparietal junction are as well. Within the DMN circuitry, a common target of these widely separated cortical mechanisms are limbic structures in the depths of the right hemisphere. Our group's discovery of the asymmetry in hippocampal causal connections now starts to be completed by the data on the analogous uniqueness in the functional connectivity of the right amygdala (Kerestes, Chase, Phillips, Ladouceur, & Eickhoff, 2017).

William James in his *Principles of Psychology* (1890) used the notion "primary memory," which was for him the equivalent of the field of consciousness — the scope of the content which we are aware of at the present moment, in the absence of physical stimulation. A bit later, Henry Bergson (1907/2006) noted that to study authentic consciousness, one has to exclude external stimulation. This reminds us of two contemporary concepts. The first is the DMN, a brain circuitry which is activated in the resting state, and is, according to the emerging knowledge, related to introspection and other aspects of consciousness. The second concept is working memory. It is a modern counterpart to James's primary memory and one of the central concepts in cognitive science because of its relevance to clinical data, intelligence measurement, and academic achievements. Recent neuroimaging studies of the DMN bring together both concepts by showing that activity around the right temporoparietal junction is the best predictor of working memory scores (Markett et al., 2017). This region in turn is the main input pathway to the right hippocampus within the DMN connectome (see Ushakov et al., 2016, Velichkovsky et al., 2017).

The ability to retain, manipulate, and invent information in the absence of external stimulation is a key prerequisite for mental life and for goal-directed interaction with the environment. However, how this ability is implemented in the brain wetware remains one of the great mysteries of psychology and neuroscience. The current collection of research papers gives us some hope that we are approaching a solution to this century-long problem.

Acknowledgments

This introduction and the majority of papers in this special issue of *Psychology in Russia* have been supported by the Russian Science Foundation (grants RScF 14-28-00234 and 17-78-30029 aimed at investigation of self-referential cognition and semantic representations in the human brain, respectively). Permanent support on the part of the journal editorial team is appreciated. Special thanks are due to T.O. Archakova, A.N. Veraksa, and Yu.P. Zinchenko.

References

Deacon, T.W. (1996). Prefrontal cortex and symbol learning: Why a brain capable of language evolved only once. In B. M. Velichkovsky & D. M. Rumbaugh (Eds), *Communicating meaning: The evolution and development of language* (pp. 103–138). Mahwah NJ: LEA Publishers. doi: 10.1111/j.1467-9744.1996.tb00954.x

Dolina, I.A., Efimova, O.I., Kildyushov, E.M., Sokolov, A.S., Khaitovich, P.E., Nedoluzhko, A.V., Sharko, F.S., & Velichkovsky, B.M. (2017). Exploring *terra incognita* of cognitive science: Lateralization of gene expression at the frontal pole of the human brain. *Psychology in Russia: State of the Art*, *10*(3), 231–247. doi: 10.11621/pir.2017.0316

Hrvoj-Mihic, B., Bienvenu, T., Stefanacci, L., Muotri, A.R., & Semendeferi, K. (2013). Evolution, development, and plasticity of the human brain: From molecules to bones. *Frontiers in Human Neuroscience, 7,* 707. doi: 10.3389/fnhum.2013.00707

James, W. (1890). *The principles of psychology*, Vols 1 and 2. New York: Henry Holt and Company. doi: 10.1037/10538-000

Kerestes, R., Chase, H.W., Phillips, M.L., Ladouceur, C.D., & Eickhoff, S.B. (2017). Multimodal evaluation of the amygdala's functional connectivity. *NeuroImag*e, *148*, 219–229. doi: 10.1016/j.neuroimage.2016.12.023

Markett, S., Reuter, M., Heeren, B., Lachmann, B., Weber, B., & Montag, C. (2017). Working memory capacity and the functional connectome — insights from resting-state fMRI and voxelwise centrality mapping. *Brain Imaging and Behavior*. doi: 10.1007/s11682-017-9688-9

Nedoluzhko, A.V., Efimova, O.I., Kildyushov, E.M., Dolina, I.A., Sokolov, A.S., Ushakov, V.L., Khaitovich, P.E. … Velichkovsky, B.M. (in preparation). Sequencing frontopolar cortex: New insights into the molecular base of human cerebral asymmetry.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.

Piaget, J. (1972). *Psychology and epistemology: Towards a theory of knowledge*. Harmondsworth, U.K.: Penguin.

Ushakov, V.L. Sharaev, M.G., Kartashov, S.I., Zavyalova, V.V., Verkhlyutov, V.M., & Velichkovsky, B.M. (2016). Dynamic Causal Modeling of hippocampal links within the human default mode network: Lateralization and computational stability of effective connections. *Frontiers in Human Neuroscience*, *10*, 528, e64466. doi: 10.3389/fnhum.2016.00528

Velichkovsky, B.M. (1999). From levels of processing to stratification of cognition: Converging evidence from three domains of research. In B.H. Challis & B.M. Velichkovsky (Eds). *Stratification in cognition and consciousness* (pp. 203–235). Amsterdam/Philadelphia: John Benjamins. doi: 10.1075/aicr.15.13vel

Velichkovsky, B.M. (2002). Heterarchy of cognition: The depths and the highs of a framework for memory research. *Memory*, *10*(5/6)*,* 405–419. doi: 10.1080/09658210244000234

Velichkovsky, B.M., Krotkova, O.A., Sharaev, M.G., & Ushakov, V.L. (2017). In search of the "I": Neuropsychology of lateralized thinking meets Dynamic Causal Modeling. *Psychology in Russia: State of the Art*, *10*(3), 7–27. doi: 10.11621/pir.2017.0301

First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society
Lomonosov Moscow State University

CONSCIOUSNESS STUDIES

In search of the "I": Neuropsychology of lateralized thinking meets Dynamic Causal Modeling

Boris M. Velichkovsky[a,b,c*], Olga A. Krotkova[d],
Maksim G. Sharaev[a], Vadim L. Ushakov[a,e]

[a] *National Research Center "Kurchatov Institute", Moscow, Russia*

[b] *Moscow Institute for Physics and Technology, Moscow, Russia*

[c] *Russian State University for the Humanities, Moscow, Russia*

[d] *N.N. Burdenko Institute of Neurosurgery, Moscow, Russia*

[e] *National Nuclear Research University "MEPhI", Moscow, Russia*

* Corresponding author. E-mail: velich@applied-cognition.org

Background. Ideas about relationships between "I", egocentric spatial orientation and the sense of bodily "Self" date back to work by classics of philosophy and psychology. Cognitive neuroscience has provided knowledge about brain areas involved in self-referential processing, such as the rostral prefrontal, temporal and parietal cortices, often active as part of the default mode network (DMN).

Objective and Method. Little is known about the contribution of inferior parietal areas to self-referential processing. Therefore, we collected observations of everyday behavior, social communication and problem solving in patients with brain lesions localized either in the left inferior parietal cortex (LIPC group, n = 45) or the right inferior parietal cortex (RIPC group, n = 58).

Results. A key characteristic of the LIPC group was an overestimation of task complexity. This led to a prolonged phase of redundant and disruptive contemplations preceding task solution. In the RIPC group, we observed disorders in reflective control and voluntary regulation of behavior. Abilities for experiencing emotions, understanding mental states, and social communication were to a great extent lost. Results are interpreted within a multilevel framework of cognitive-affective organization (Velichkovsky, 2002). In particular, we highlight the role of right-hemisphere mechanisms in self-referential cognition, emotional and corporeal awareness. This is consistent with recent data on a profound asymmetry in connectivity of left and right hippocampi within the DMN (Ushakov et al., 2016).

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0301
http://psychologyinrussia.com

Conclusion. It seems that the center of egocentric spatial representation plays a special role in accessing self-related data. Normally, the right hippocampus provides a holistic representation of surrounding and, thus, an easy-to-find gateway into much of what we used to call "subjective experience". This heuristics becomes misleading in the case of right-sided brain lesions.

Keywords*:* thinking, emotions, lateralization, hippocampal formation, neuropsychology, dynamic causal modeling (DCM), egocentric spatial orientation, Self-referential cognition, levels of cognitive organization

In Memoriam: Bruce Bridgeman, 1944-2016

Introduction

Thinking in patients with brain damages of different etiology remains a relatively weakly studied chapter of cognitive neuropsychology, both from point of view of diagnostics and that of rehabilitation. Despite a substantial number of diagnostic tests and an abundance of fractional data, the overall picture of this central neurocognitive issue still is fragmented and contradictory. The lack of a conceptual *Gestalt* makes it difficult to elucidate factors influencing learnability and find the ways to a better social adaptation of patients. Classical research devoted to the analysis of cognitive impairments in solution of arithmetic tasks (Luria, & Tsvetkova, 1966), verbal-logical inferences (Balonov, Deglin, & Tschernigovskaja, 1979), or visual-spatial constructive tests (Khomskaja, 1987; Korsakova, & Moskovichute, 1988) paid relatively little attention to the everyday problems of patients, though such problems are a particularly importance source of data for cognitive conceptualization. With a few exceptions, studies of the modern neuroimaging era demonstrate even less interest to the modeling of everyday tasks and situations (see, e.g., Gazzaniga, 2009).

In the present article, we attempted to combine a more phenomenological approach, which takes into account everyday problems of patients suffering from unilateral brain lesions, with new knowledge about human brain structural, functional and in particular effective (cause-and-effect) connectivity. In one such development, a recent diffusion tensor imaging (DTI) analysis of anatomical white matter asymmetries across the whole brain of 41 children and adolescents with Autism Spectrum Disorders (ASD) and a matched control group of 44 typically developing (TD) participants revealed that children with ASD have reduced lateralization compared to TD children who showed significant asymmetry with rightward anisotropy (Carper, Treiber, DeJesus, & Müller, 2016). These findings can be interpreted as reflecting different processing modes in two hemispheres. The "division of labor" between hemispheres appears to be diminished in ASD, possibly underlying the characteristic pattern of this group's deficiency in social intelligence.

In another line of research, systematic hemispheric differences in molecular mechanisms were discovered even for closely located brain regions, such as the frontopolar Brodmann Areas 10 on the left (BA10L) and on the right (BA10R). Similarly, this research shows that most of the strongly expressed genes — and almost all of the differentially expressed protein-encoding genes -- were detected in the right frontopolar cortex (Dolina et al., 2017). A neuropsychological *pendant* to these data is the well-established knowledge that if the left hemisphere supports basic linguistic functions, the right prefrontal cortex might be important for understanding of metaphorical language, humor, irony and sarcasm (Balonov, Deglin, & Tschernigovskaja, 1985; Krotkova, & Velichkovsky, 2008; Shammi, & Stuss, 1999). In addition, right prefrontal areas are mainly involved in processes of autobiographical memory and personal planning for the future (Dickerson, & Eichenbaum, 2010). Both these groups of processes are directly related to our subjective experience, i.e., to our conscious «Self».

The picture of prefrontal (likely right-sided) involvement in the higher-order metacognitive processing is appealing and receiving support (Craik et al., 1999; Sokolov, 2013; Stuss, Rosenbaum, Malcom, Christiana, & Keenan, 2005; Velichkovsky, Klemm, Dettmar, & Volke, 1996) but it may be incomplete. There is another region seemingly realizing similar functions with respect to self-referential cognition, e.g. as related to retrieval of autobiographical memories and personal planning for the future. The region includes the inferior parietal lobe and the temporoparietal junction. Thereafter, we will call these 'left and right inferior parietal cortex' (LIPC and RIPC, respectively). Basically, these are structures of higher-order multimodal sensory integration, whereby LIPS is responsible for representation of the right and RIPS for representation of the left side of the surrounding.

One reason for the increased interest in this region is its role in the default mode network (DMN), a set of interconnected brain areas that are activated in the resting state and deactivated by any cognitively effortful task (Arsalidou, Pascual-Leone, Johnson, Morris, & Taylor, 2013; Gusnard, Akbudak, Shulman, & Raichle, 2001; Raichle et al., 2001). Hypotheses on the DMN functionality have also been formulated mostly relating it to higher-order aspects of consciousness and cognition (Gusnard, & Raichle, 2001; Schilbach, Eickhoff, Rotarska-Jagiela, Fink, & Vogeley, 2008). The main parts of the DMN have been identified in the medial prefrontal cortex (MPFC), posterior cingulate cortex (PCC), and the inferior parietal cortex of both hemispheres (Buckner, Andrews-Hanna, & Schacter, 2008). With respect to hippocampal formation, rank correlations of activity also reveal the pattern of activation/deactivation characteristic of the DMN (Vincent, Bloomer, Hinson, & Bergmann, 2006). Connectivity patterns of hippocampal formation are of particular interest because of its crucial role in episodic memory processes (Dickerson, & Eichenbaum, 2010) and in representation of surrounding space (Burgess, Jackson, Hartley, & O'Keefe, 2000; Moser, & Moser, 2008).

Functional connectivity of both hippocampi has been analyzed extensively, for instance, in a recent meta-analytic study by Robinson, Salibi and Deshpande (2016). However, functional data have a low scientific status, as they are only correlational in nature. Therefore, of importance are studies where effective (cause-and-effect) relations among left and right hippocampal formation (LHIP and RHIP, respectively) and other DMN structures have been for the first time investigated by a combination of the functional magnetic resonance imaging (fMRI) and the mathematical method of spectral dynamic causal modeling (DCM). The method's main idea is to evaluate parameters of a biologically-validated model of the neuronal system so that it could predict the observed fMRI data in the best way (Sharaev, Zavyalova, Ushakov, Kartashov, & Velichkovsky, 2016; Ushakov et al., 2016). These studies conducted on a group of 30 healthy right-handed subjects led us to the discovery of a profound asymmetry in LHIP and RHIP effective connections. Fig.1 illustrates this asymmetric pattern of interactions. LHIP demonstrated a high involvement in the DMN activity, with information outflow preponderant to all other DMN regions including RHIP, as shown by our analysis of two 5-nodes and one 6-nodes interactions. Causal interactions of LHIP with inferior parietal cortex were bidirectional only in the case of LIPC: there was not inflow to LHIP from RIPC. This means that in terms of spatial representation LHIP had access to information only about contralateral, right hemispace. On the contrary, RHIP was affected by inputs from both LIPC and RIPC that would allow a holistic — left and right-sided — multimodal representation of egocentric space (for a detailed analysis of the models, see Ushakov et al., 2016).

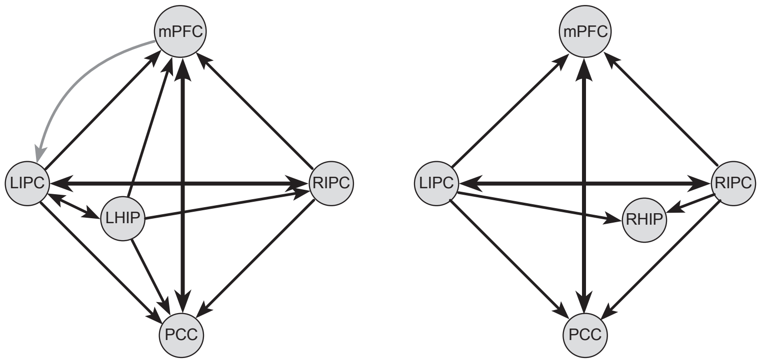


Figure 1. Winning models of effective connectivity patterns for left and right hippocampal formation on the left and the right side of the figure, respectively. The connection between medial prefrontal cortex (mPFC) and LIPC on the left side is inhibitory. (After Ushakov et al., 2016, with permission of the Publisher — Frontiers Research Foundation)

In our view, this pattern of asymmetry in effective connections of the hippocampal regions may be related to lateralization phenomena in verbal and spatial domains known in human neurophysiology, neuropsychology, and neurolinguistics.[1] As a matter of fact, there is an obvious drawback of such lateralized architecture: a destruction of RHIP or RIPC could lead to the left-sided spatial hemi-neglect not compensated by preserved LHIP/LIPC interconnections. This phenomenon is well-known from clinical data (Harrison, 2015; Howard, & Templeton, 1966; Luria, 1966). Distortions of corporeal awareness such as out-of-body experience (Blanke, & Mohr, 2005), asomatognosia (Baier, & Karnath, 2008) and anosognosia (Heilman, 2014; Vallar, Bottini, & Sterzi, 2003) have also been described with the same locus of lesions in the posterior part of the right hemisphere. Lateralization of higher-order cognitive and emotional processes in patients with local brain damages to either of LIPC or RIPC is by far less investigated though it could be supposed in light of previous observations (Krotkova, & Velichkovsky, 2008; Singh-Curry, &

[1] In mice, there is evidence on differences in synaptic plasticity and long-term potentiation between the left and right hippocampi (El-Gaby, Shipton, & Paulsen, 2015; Shipton et al., 2014). No asymmetry in connectivity of hippocampal formation has been reported in animal studies (Edvard Moser, personal communication to the first author, April 13, 2016).

Husain, 2009). Comparative analysis of these processes in the relevant groups of patients was the primary objective of this study.

Method

Clinical material

We based this study on descriptions including more than 270 clinical cases of patients with unilateral brain damages of different etiology (e.g., traumata, tumors and blood vessel dysfunctions). Over several years, all the patients were observed in their everyday behavior and went through neuropsychological intervention programs in the division of rehabilitation at the Institute of Neurosurgery named after academician N.N. Burdenko in Moscow. From this database, we excluded cases in the following two categories: first, patients with relatively narrow lesions of primarily and secondary sensory regions (according to well-known neurological criteria — see Kolb, & Whishaw, 2015; Luria, 1966); second, patients with substantial damages of prefrontal regions. These latter damages would create specific difficulties for neuropsychological analysis due to diverse manifestations of dysexecutive ("frontal-lobe") syndrome such as instability of attention, lack of motivation and general inactivity. The main reason for the exclusion was however our intention to select a target group of patients with lesions in the posterior tertiary regions of cortex overlapping with the loci of LIPC and RIPC to be consistent with data on effective connections within DMN network.

As a result, two samples of patients with corresponding localization of brain damages were selected for a close analysis of their everyday behavior: 45 cases (24 females), mean age 48 years, with left hemisphere localization and 58 cases (31 females), mean age 45 years, with lesions in the right hemisphere, all around inferior parietal lobe and temporoparietal junction. According to well-known classification of Brodmann cytoarchetechtonic maps (e.g. Kolb, & Whishaw, 2015), lesions included BA 39. Often, neighboring areas BA22, 37, and 40 were also part of the lesion. This localization, which roughly corresponded to the definition of LIPC and RIPC in the DMN studies, was confirmed by structural MRI imaging and at times by the data of neurosurgical interventions. In what follows, we consider the results of our phenomenological analysis for the LIPC and RIPC groups, in that order.

Results

Phenomenology of the left hemisphere lesions

The central phenomenon in the behavior of LIPC patients was observed in a wide range of situations. We call it the "difficulty of entrance in the task" (DET), for the sake of simplicity. For example, during rehabilitation-exercises, a typical observation was that patients — after attentively following instruction to an exercise — started to work on it so poorly that this lead to an impression of their total inability to solve such category of tasks. However, after some delay filled with detailed explanation of the task by the neuropsychologist, patients suddenly arrived at the solution fast and correctly, leaving open the possible reason for their initial problems. The DET phenomenon is quite different to the dysexecutive ("frontal-lobe") strategy of behavior, which is defined by impulsivity in decision making, as well as instability of attention and motivation. Here, in contrast, the level of achievement motivation was constantly elevated with often rather good scores in attention and immediate memory tests. Obviously, the difficulty was somehow related to understanding of the instruction and interpretation of the task situation by patients.

The same DET pattern of behavior was seen practically in all neuropsychological probes. We observed it even in very easy tasks that are typically clear from the scratch, without any additional explanations. For instance, in the Seguin Form Board test, where one has to place simple wooden forms into corresponding holes of the board, there usually is no need of instruction. Most subjects proceed to the correct solution in response to simply an inviting gesture and an encouraging head nod. This was not the case with LIPC patients. They started to explore forms, to sort and lay them out in a row etc., performing manipulations in no way related to the obvious solution. At the same time, there was no problem, in principle, in reaching the solution: shapes of the blocks were perfectly recognized and no difficulties in eye to hand coordination was observed. The problem was in the interpretation of the task situation. After a successful solution, if patients were asked for the reason of their initial reluctances, a typical answer was that at the beginning his/her first impression was that one had to find something complex and previously unknown which could not be immediately clear from the situation. Therefore, a straightforward solution was ignored.

To our knowledge, the DET phenomenon has never been stated in such a form in the past though there were many reports on categorization problems in patients with damages of left temporal lobe and temporo-parieto-occipital junction (Gelb, & Goldstein, 1920; Koivisto, & Laine, 2000; Wilkins, & Moscovich, 1978). As a matter of fact, a typical neuropsychological investigation focuses on analysis of task solution *per se*, i.e. the analysis starts only after the instruction is completely understood by the subject. This is possibly why the DET phenomenon, which is in essence the unusual interpretation of task situation, remained without due attention.

To model DET phenomenon in a common neuropsychological context, we sampled contrasting groups of patients with lesions in LIPC (n = 9) and RIPC (n = 8). Participants were given the Wisconsin Card Sorting Test (WCST). Test instruction is open to many interpretations: "I cannot explain to you how to solve this task. Please take cards one by one from the pile and place them on the four keycards. I will tell you every time whether you did it correct or incorrect". On the cards there are figures which differ in shape, color and number. Subjects lay down cards by trial and error receiving experimenter's evaluation as "correct" or "incorrect". The category is considered to be learned if there are 10 errorless trials in a row. After that, the experimenter normally changes the categorization rule and the whole procedure repeats until the next 10 correct trials occur. In our study, we concentrated our analysis on the very first classifications registering the number of sortings needed to confirm understanding of the task instruction.

This reduced version of WCST revealed dramatic differences in performance of both patients' groups. The RIPC group members needed only 2.1 trials on average to figure out the principle of the task and come up with a definite strategy of solution, whereas the LIPC patients needed as much as 12.2 trials to arrive at this understanding. These differences were highly significant (Wilcoxon-Mann-Whitney rank sum test, $U=9$, $p<0.01$). In should be stressed that following test classifications, when the first 10 errorless trials were achieved, differences in performance of both groups became non-significant. Similarly, no significant differences were observed between LIPC and RIPC patients in additional tests of working memory, selective attention and task switching. In fact, they often demonstrated better scores than "right-sided" patients. As to aphasic disturbances in LIPC patients, those were not serious enough to prohibit a dialog with experimenter. Thus, one can conclude that the DET phenomenon was successfully modeled in the experiment: our LIPC and not RIPC patients demonstrated selective problems in the initial phase of task situation by ignoring the simple and easy available solution. Of particular interest are self-reports of LIPC patients explaining their difficulties in finding a solution. Let us take color as the critical category. Typically, subjects with intact brains may make one or two wrong selections (shape or/and number) but thereafter come to the correct guess and solve the task. In LIPC patients, this expected sequence of events was never observed. One patient (female, 27 years, higher education, lesion resulted from a gunshot, with an alien object in the posterior parietal parts of left hemisphere), could not reach the solution after 25 trials. This happened despite her efforts to work with high accuracy. We interrupted the session in view of patient's strain and negative emotional reactions to "incorrect" remarks. When asked to name features of cards, she mentioned shape, number, structure, spatial configuration but not color. Then we asked "And what about color?" — "Yes, color too, here is red, there are green, blue and yellow". — "Why did you never attempt to sort cards by color?" — "I thought this would be too easy for a solution and that you gave me a more complex task". Even if other members of the same group came up with color as the relevant category, they continued treating the task as being more complex than it in fact was. In three such cases we interrupted the test after 20 trials without solution. Afterward, these patients said that the emphasis on color seemed to be insufficient for them, so they looked for some "sequence algorithm" in the sorting. We also observed rather unusual hypothesis, when, for example, one patient decided to match cards to key samples in such a way that "no features will be in common". Of course, such an excessively reflective strategy could not be successful in the WCST. All attempts at solution were done with maximum of efforts; sometimes one could see tears in patients' eyes after next unsuccessful trial.

Thus, neither an attention deficit, nor memory weakness, nor lack of achievement motivation can directly explain this specific difficulty we detected on the early stage of task solution in patients with damage in left posterior cortical areas. Their problem lies in a general attitude towards task situations as *a priori* unique, i.e. requiring tough mental efforts and sophisticated strategies of problem solving. In the case of neuropsychological tests as well as everyday task situations, which all have low or middle levels of complexity, such a mode of thinking leads to overseeing obvious ways towards solution. "Permanent misunderstandings" were reported by relatives of these patients. We also observed these particularities during reha-exercises when simple movements could present LIPC patients with insurmountable strain, which led them to be suddenly "frozen" in an astonishment posture. The only way of overcoming such episodes was to start the explanation anew in an explicit top-down fashion: "We will now learn to walk properly, for this we have to perform several easy to learn exercises... To start with please repeat the movement, which I am showing you now". There were large individual differences in the revealed picture of the "left-sided" mode of thinking which was observed on background of aphasic and motor disturbances to a different degree typical for patients. However, the central phenomenon of overcomplicating any task situation, as if it would need particular cognitive efforts, was present in all LIPC patients.

Phenomenology of the right hemisphere lesions

A completely different set of difficulties was observed in our RIPC patients. Before their description, we wish overview some peculiar features in behavior of these patients known from the literature and confirmed by our observations (see, e.g., Balonov, Deglin, & Tschernigovskaja, 1985; Kolb, & Whishaw, 2015; Luria, 1966). Right-handed RIPC patients do not usually have aphasic disturbances: their speech remains intact both grammatically and lexically. However, one often finds deficits in speech intonation structure. The patient's voice loses its normal modulation of volume and cannot be voluntary regulated in a socially appropriate manner: it is either too quiet or loud and crude. Prosody of speech is also distorted — it becomes voiceless, husky, nasal or barking and shrill. Spontaneous speech makes a strange impression as monotonous and having no emotional expression. This cannot be changed even after an explicit instruction by the neuropsychologist, so special exercises are needed to correct the deficit. As a rule, this poverty of expression coexists with similar deficits on the side of speech perception seriously complicating interpersonal communication (Koelsch, Kasper, Sammler, Schulze, Gunter, & Friederici, 2004). For instance, if one reads to the patient the same phrase with three intonations, that of doubt, mockery or fright, he/she will be unable to distinguish the variants by hearing.

Most of emotional expressivity is also lost in facial expression and gesture. The face loses its liveliness and gaze seems to be "frozen". Often it is difficult for patients to recognize themselves in old photos "before" the trauma, as if it were other persons. Facial reactions do not disappear completely but e.g., smile has an unnatural, torturous character or becomes a coloration of euphoric comfort and some stupidity. It should be noted that patients with damage of left hemisphere demonstrate an opposite pattern of communicative abilities. Even with strong language disturbances like the aphasia of Wernike type, when patients have no single correct word in their lexicon so speech is a "word salad", we can understand almost everything that they wish to say or to ask, can feel their mood and often maintain a rather informative dialog thanks to their intact facial expression, gesture and intonation.

With respect to problem solving behavior, RIPC patients do not demonstrate the slightest signs of DEP phenomenon. They easily manage the WCST and the variety of neuropsychological tests "on thinking". However, a seemingly simple task from another domain suddenly is difficult for RIPC patients. For example, when we present them a picture depicting several personages who are in definite social relations to each other and the task is to explain these relations by thinking out loud one or two phrases for each personage, i.e. what everyone of the personages could say in this situation. Un our practice, we use for such tasks from "Stories in pictures" by N. Radlov (Radlov, Harms, Dilatorskaya, & Gernet, 2015/1937). This is a comic book for children in senior kindergarten. Though the pictures are without captions, depicted situations are so clear that children delightfully invent dialogs as all narratives have some humorous gist.

This sort of tasks led to substantial difficulties in our RIPC patients with lesions. They were able to –in general terms– describe the situation but could not reconstruct moods, intentions and possible remarks of depicted personages. As a rule, the concealed meaning, the very humorous spirit of every story in pictures was not discovered. Here is one example. In a series of three pictures, a man is depicted walking in a park with two puppies. When the wind takes his hat off, he orders puppies, by an imperious gesture, to bring it back. They do but tear the hat apart straggling for the master's favor. A third picture shows the man whose posture and facial expression testify to his distress and confusion. After a long examination of pictures, one of the patients described the story in the following way: "A man walked with dogs. He threw them two hats that they brought back. The brown dog returned first and was praised by the man". Here is another example which includes two pictures. On the first, there is a boy who undresses to go swim in a river and, without looking, he puts his hat on the horn of a cow standing in the bush behind him. The second picture shows the cow walking with the hat on the head, as well as the boy and a calf, both observing this with extreme astonishment. The patient's story was as plain as following: "The boy hung his hat on the cow's horn and she went away". There were no appropriate descriptions of internal states, emotional exclamations or hypotheses about social interactions in the situation. Similar problems were detected with all forms of representations, including familiar photos and paintings. Being confronted with a reproduction of N. Ge's "Peter I interrogating the Tsarevich Alexei in Peterhof" another patient correctly recognized the painting and remembered some events preceding this scene. He also easily described inanimate details of the interior. However being asked about feelings of persons in the scene, he replied in an overtly inadequate manner "Peter is in a good mood, he is joyful. Alexei feels conceit".

Deficiencies in understanding emotional aspects of observed interpersonal communication coincided in RIPC patients with deficits of self-consciousness extended to their-own affective and mental states. Being presented with a set of photos showing people with different emotional expressions, they cannot find which one reflected their feelings and mood at that moment. For instant, a patient, in a state of intense irritation and just after two aggressive attacks against his mother sitting nearby, selected as a descriptor of his state the photo of a boy whose smiling face almost "eradiated" happiness. After a long search for an appropriate photo, another patient said: "No, there is no such photo here as I do not feel anything". What we have seen as a dominant mood in these patients is a neutral placidity; sometimes it is interrupted by bursts of irritated aggression but there are no episodes of worries, fear or vivid happiness. Along with emotions, they seemed not to express common states such as fatigue. One of them, while obviously exhausted, negatively responded to our question about possible fatigue. Still another patient formulated his sensations in the following way: "I always cannot understand what it means "to be fatigued". Should I have some pains? But I have no pains, nowhere".

Besides these specific problems of self-referential cognition, we observed weakening of voluntary control of behaviors and meta-cognition in our RIPC patients.

We already noted their difficulties with voluntary regulation of voice and emotional control. But this weakening had a more global character and is not necessarily focused on the emotional sphere. A patient could easily follow instruction "Close your eyes" but if you asked him to do something else, e.g., to grope about an object, the eyes reopened involuntary. A reminder that eyes should be closed did not work though patient was perfectly inclined to follow it: as soon as attention switched to another task, the eyes reopened to the astonishment of the patient. The same pattern applies to "asocial movements", observed during patients' involvement in complex tasks. A patient with a high premorbid status, adequate in social life and successful in almost all tests of the neuropsychological examination, could suddenly start scratching his body, digging into his nose or picking on something in his hairs, when a task demanded his full attention. After a remark from the experimenter that this is inappropriate in social situation, he got confused for a while but reappeared the behaviors if the task solution was effortful enough.

As to meta-cognition, i.e. deliberate regulation of own mental processes, it is weakened to the same substantial degree as control over external behaviors in RIPC patients. If we have nothing to do for some period of time, we are normally engaged in experiencing a kind of William James' stream of consciousness which is in part under our control because it is always possible to turn its direction either to planning for the future or to remembering images from the past vacation. Our RIPC patients claimed that when they, for example, waited outside the room to begin their reha-exercises they had no thoughts. In order to launch thinking process, they seemed to need an external stimulus such as the advice of another person.[1] There were no complains about the lack of thoughts. Apparently, the patients did not experience boredom and had no intention to entertain them-selves in any way.

The emptiness of mental life was extended to self-referential aspects of remembering. Though patients' performance in ordinary memory tests could be nearly perfect, we learned about salient autobiographical events of their lives almost exclusively from relatives because, as a rule, patients were unable to remember such subjectively colored information. Even photos of an event were of little help: only formal knowledge such as names of participating persons and general circumstances was retrieved but nothing that was mediated by subjective experience, either his/her-own (first-person perspective) or participants (second-person perspective) represented and correctly recognized in the photo. By borrowing terminology from A.N. Leontiev's activity theory, only meaning was remembered not personal sense.

As a whole, these phenomena build a coherent picture which testifies to disrupted comprehension of interpersonal relationships in our RIPC patients. It clearly dissociates with their relatively intact formal knowledge and cognitive operations on information about inanimate objects, i.e. the domain of thinking in which it is possible to ignore subjectivity: feelings, beliefs and intentions. For the patients and their relatives this is a serious factor of invalidation in everyday situations. Errors with pragmatic context of communication, misunderstanding of other people's

[1] This is similar to dysexecutive syndrome usually related to damage of prefrontal regions (*aka* "frontal lobe" syndrome). In particular, problems with multitasking were previously related to damage of frontopolar cortex (Penfield, and Evans, 1935; Burgess, Cohen-Yaacovi, & Volle, 2012).

emotions and inability to empathy destroy their social life: the mutual understanding with relatives disappears, friends start avoiding contacts, and overall alienation grows. In direct communication with others, they demonstrate symptoms of autism spectrum disorder, also with respect to their aberrant eye movement behavior eluding visual fixations on face and in particular on eyes of their *vis-à-vis*. Not less complicated is any form of indirect communication. For example, speaking to such a patient on the telephone, you never know for sure whether he/she is interested in a conversation or will hang up the next moment.

Discussion

We describe two qualitatively different patterns in everyday behavior and test performance in patients with damage in either left or right tertiary areas in the posterior cortex. Although there are large individual differences and interfering influences of other deficits (e.g., speaking and movement control in the LIPC group), the two patterns are very distinct. LIPC patients tend to overestimate the actual complexity of their surrounding world of things and standard social situations (such as in neuropsychological testing), whereas RIPC patients simplify complex social interactions by failing to attribute mental states to other people and experience these states by them-selves suggestive of deficits in self-referential and interpersonal cognition[1]. There are practical and theoretical consequences of these distinct characteristics. From the practical point of view, our results advocate for differential procedures in neurorehabilitation of patients with LIPC and RIPC damage. As even these procedures are not theory-free, a theoretical explanation of the emerging picture of a cognitive-affective architecture is the task of foremost importance, both for conceptual development and for practical applications. It should be said that the discovered pattern of hemispheric asymmetries questions the validity of traditional dichotomies and demands for a re-engineering of existing approaches.

The first of these dichotomies considers the left hemisphere as "leading" (or "dominant", with respect to performance in simple sensorimotor probes in response to verbal instruction). This old dichotomy evidently reverses the real interactions and relative importance of lateralized brain mechanisms. In fact, damage of right hemisphere has a more serious deteriorating effect on everyday activities and social life of patients. This, in a sense, testifies to the leading role of the right hemisphere in realization of specifically human daily tasks. Another classical distinction of hemispheric asymmetry is that of "left is verbal and right is nonverbal", which does not fare much better. Firstly, it is nonspecific, particularly with respect to the role of right hemisphere. Secondly, it ignores involvement of right hemisphere mechanisms in different forms of verbal processes (Federmeier, Mai, & Kutas, 2005; Winner, & Gardner, 1977), as well as the considerable role of left hemisphere in spatial operations of mental rotation, translation and zooming of imagined objects (Cohen et al., 1996; Mehta, & Newcombe, 1991). Thirdly, this distinction does not correspond to the rather not trivial picture of deficits revealed by the present study.

1 In the context of early development, a strong relationship of both forms of cognition — self-related and other-related — has been shown by Doris Bischof-Koehler (Bischof-Koehler, 1989).

A standard model of hemispheric asymmetry popular in the last two decades explains it in terms of differential functioning in semantic memory (Atchley, Story, & Buchanan, 2001; Burgess, & Lund, 1998; Chiarello, 1998). Accordingly, the left hemisphere supports retrieval of high frequency associations, whereas the right hemisphere supports activates semantic relations with low frequencies of previous use. The model has been mainly used in neurolinguistics for explaining data on the hemispheric asymmetry in understanding of metaphoric language and indirect speech acts (Balonov, Deglin, & Tschernigovskaja, 1985; Shammi, & Stuss, 1999; Winner, & Gardner, 1977). Even in this narrow domain, the view does not receive support in more recent neuroimaging (Forgács et al., 2012) and divided visual field (Forgács, Lukács, & Pléh, 2014) experiments. The main problem with this hypothesis is that the observed phenomenology of hemispheric differences is richer than one implied by the frequency of semantic associations. Moreover, the theoretically implied direction of difference is opposite to one observed in reality. For example, the model cannot explain why damage of right hemisphere could lead to selective deficits in self-referential cognition and theory of mind, i.e. patients' knowledge about knowledge, emotions and intentions of other people. Obviously, thoughts about emotions and intentions of one-self and others are especially frequent in the mental life of a typical healthy person.

In experiments, cognitive tasks, from perception of form to encoding information in terms of personal sense, show different patterns of asymmetries. For example, evaluation of some material as belonging to a certain semantic category leads to primary activation of left hemisphere, whereas encoding the same material in terms of its personal sense for the subject mostly activates right prefrontal cortex (Velichkovsky, Klemm, Dettmar, & Volke, 1996). In addition, hemispheric differences seem to have a long evolutionary history (Karenina, Giljov, Ingram, Rowntree, & Malashichev, 2017; Ocklenburg, & Güntürkün, 2012), therefore a broader approach describing several evolutionary steps, or levels in cognitive-affective organization need to be considered. Presuming that there is such a "vertical dimension" of mental functioning, what could granularity and distinct characteristics of levels be? It is clear that dichotomies are too unspecific. In the same vein, disagreement between authors of three-level theories implies that more levels may be at work[1] . The founder of biomechanics, N.A. Bernstein (1947), described four levels, from A to D, involved in realization of human movements. One of us upgraded his views some time ago, which led to a Grand design model with as many as six different levels of organization (Velichkovsky, 1990). The first group (from A to D) is primarily built up by the sensorimotor mechanisms. The second group (from E to F) consists of mechanisms of higher symbolic coordination. Here is the list of these levels in bottom-up order in a version which is about 15 years old (Velichkovsky, 2002, pp. 406–407).

> <u>Level A: Paleokinetic Regulations</u>. Bernstein also called it the "rubro-spinal" level, having in mind the structures of spinal cord and brain stem (up to midbrain) involved

[1] In the history of psychology and cognitive science, many authors — as different as Wundt, Vygotsky, and Fodor — used two-level theoretical constructs in their work. Well-known three-level theories were developed by Karl Buehler, Jens Rasmussen, and, in human neuroscience, by John Houghling-Jackson, A.R. Luria and Paul MacLean (see Velichkovsky, 2006).

in regulation of the muscles' tonus as well as paleovestibular and basic defensive reflexes. The awareness of functioning is reduced here protopathic sensitivity (Head, 1920), which is so hedonistic, diffuse, and lacking any precise spatial coordinates (any definite "local signs") that even the term sensation seems to be too intellectual in this case.

Level B: Synergies. Due to evolution of new neurological mechanisms—the "thalamo-pallidar system" after Bernstein—the broad sensory integration and regulation of the organism's movements as a whole become possible, transforming it into a "locomotory machine". The specializations of this level are movements involving large groups of muscles of different body parts, e.g., rhythmic and cyclic patterns of motion underlying all forms of locomotion. Possibilities of awareness are limited to proprio- and tangoreceptoric sensations within the body's frame of reference.

Level C: Spatial Field. The next round of evolution adds exteroception to the repertoire of sensory corrections. This opens outer 3d space and makes possible one-time goal/place-directed movements as well as topographically contingent behavior in the near environment. The control instances of the level are phylogenetically new parts of basal ganglia (striatum) and stimulotopically organized cortical areas, especially in posterior parietal cortex. The corresponding subjective experience is that of a stable voluminous surrounding filled with localized but only globally sketched objects.

Level D: Object Actions. A new spiral of evolution leads to the building of a variety of secondary areas of neocortex with parietal, premotor, and partially temporal regions as the main instances. This permits detailed form perception and object-adjusted manipulations. Individualized objects affording some but not other actions come to the focus of attention. Formation and tuning of sophisticated higher-order sensorimotor and perceptual skills is supported by a memory of the procedural type. Phenomenal experience is the perceptual image (as described by Gestalt school — e.g., Koffka, 1935).

Level E: Conceptual Structures. Supramodal associative cortices of temporo-parietal and frontal structures, particularly on the left side, provide the highest integration of various modalities supporting the ability to categorize objects and events as members of generic classes. Development of language and culture fosters this ability and virtually leads to formation of the powerful declarative-procedural mechanisms for symbolic representation of knowledge (widely but not quite correctly known as semantic memory). Common consciousness is the awareness mode at this level.

Level F: Metacognitive Coordinations. Changes in conceptual structures result not only from accretion of factual experience but also from experimentation with ontological (truth-value) parameters of knowledge. This "personal view of the world" and its counterpart, "theory of mind", are supported by those parts of the neocortex that show largest growth in anthropogenesis, notably by the prefrontal, especially, right prefrontal regions. This level provides resources for dealing with novel situations and tasks without (known) solution. It is behind self-referential and interpersonal processing, reflective consciousness, and productive imagination.

In the last decade, most of our experimental efforts aimed at refinement of the Grand design approach have been focused on two middle levels, C and D. These levels were related to both major pathways in development of sensory systems, dorsal and ventral "streams" (Velichkovsky, 2007; Velichkovsky, Joos, Helmert, & Pannasch, 2005). Seminal research on the role of hippocampus in episodic memory and in representation of surrounding space (Dickerson, & Eichenbaum, 2010; Moser, & Moser, 2008) opened the way to understanding of respective integration mechanisms in paleocortex whereby, in primates, dorsal stream information ("Where?") propagates via parahippocampal structures and medial entorhinal area while ventral stream information ("What?") seems to access the hippocampal formation through entorhinal cortex and lateral entorhinal area. Our recent data on effective connectivity of both hippocampi (Ushakov et al., 2016) show that a holistic multimodal representation of the surrounding space can be achieved only by the right hippocampus (see RHIP in Fig.1). This new result suggests a leading role of the right hemisphere with respect to primarily tasks of Level C. As to the present study, its main contribution is in correcting previous views ascribing metacognitive functions solely to prefrontal regions. Clearly, damages to the region including inferior parietal lobe and temporoparietal junction, namely RIPC, result in the kind of disturbances which could be expected after removal of mechanisms responsible for self-referential and interpersonal processing (Level F, of the Grand Design model).

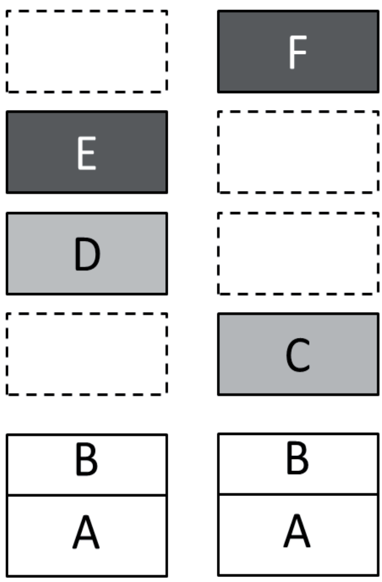


Figure 2. Grand Design model with the presumed asymmetry of multilevel mechanisms in left and right hemispheres. Empty boxes mean that some levels are, at least, underrepresented on the side. (See text for explanation)

Let us illustrate this by means of a scheme in Fig.2, where left and right sides signify structures of the left and right hemispheres. A removal of the upper box on the right side (i.e. the dark box with "F" on it) would lead to consequences which are simultaneously dramatic and very simple: the system loses its highest level of organization and as a result demonstrates the whole spectrum of disorders in reflective control and voluntary regulation of behavior. These negative changes are especially salient in the case of interpersonal relations and self-consciousness. Despite relatively preserved language mechanisms and intact basic cultural knowledge, patients with lesions of right hemisphere to a great extent lose their abilities for social communication. Moreover, together with emotional experience most of their personality vanishes as well. A completely different pattern of disturbances would arise after a removal of upper structure in the left part of Fig.2 (the dark box with "E"). This would lead to an unusual misbalance of the system's architecture. The outfall of Level E (Conceptual Structures) with survival of Level F (Metacognitive Coordinations) would result in a paradox tendency of interpreting every task situation, even a trivial one, as a new challenge demanding some creative efforts. One can easily recognize the DET phenomenon in such surplus of redundant and often destructive contemplations. This phenomenon was systematically observed in problem solving behavior of our LIPC patients.

Thus, we demonstrated how differences in thinking of patients with disturbances of left and right hemispheres could be explained within the unified Grand Design framework. Of importance is however a more detailed understanding of relations between the levels with respect to their brain mechanisms. Large-scale brain mechanisms are most of the time in a state of dynamic balance. Any negative change in the architecture leads to a number of transformations, sometimes even to an exaggerated growth in other domains as it is the case with language development of children with Williams syndrome who otherwise have substantial deficits in spatial perception and thinking (Bellugi, Lichtenberger, Jones, Lai, & St. George, 2000). In our discussion of LIPC and RIPC patients, we emphasized symptoms characteristic for their specific modes of thinking after outfall of higher symbolic coordinations, but they also have a number of disorders in other neurological domains. In Fig. 2, main loci of the concomitant problems are marked with gray color. Damages to tertiary regions of the left hemisphere are frequently accompanied by pareses and dyspraxia as well as by disorders of object perception and reading. Within the Grand Design framework, these problems can be localized on the Level D.

A particularly variegated set of deficiencies can be observed in RIPC patients. Up to one third of them demonstrate neglect phenomena in one form or another. Most often this is the classical left-sided spatial hemi-neglect (Howard, & Templeton, 1966; Luria, 1966), which is a marked disorder of spatial field mechanisms, or Level C in the Grand Design model. But there are also distortions of corporeal awareness such as out-of-body experience (Blanke, & Mohr, 2005), asomatognosia (Baier, & Karnath, 2008) and anosognosia (Heilman, 2014), which are more difficult to attribute to spatial perception. In fact, this combination of symptoms ranging from distortions of bodily Self ("*Koerper Ich*" of old German authors) to that of higher-order thought, social intelligence and emotional processes is one of the greatest riddles in neuropsychology and cognitive science. Notions such as 'embodied cognition' are of not much help as they only rename the problem in unspecific terms.

Our working hypothesis is this. The role of *explanans* has the newly discovered right-ward lateralization of spatial representation abilities in the human parahippocampal regions (Ushakov et al., 2016). As we emphasized it above, cause-and-effect connections of the left –in contrast to the right-- hippocampal formation do not allow for a holistic representation of the surrounding space. If RIPC, its input to RHIP or perhaps RHIP itself are damaged, then all the tasks demanding a kind of personal appraisal may become problematical as an access to self-related cognitive-affective data cannot be easily found. Indeed, where could data related to "Self" be most easily found in the brain? As brain mechanisms have neither time not abilities to consult philosophical dissertations, it should be a simple heuristic. The simplest one is to search for self-related data at the obvious "Self" location, i.e. at the center of egocentric spatial representation. In normal conditions, it is the right hippocampal formation which provides the easy-to-find gateway into much of what we used to call "our subjective experience". However, after right-hemisphere lesions, representation of "Self" may disappear or be somewhere shifted and lost within the scrambled spatial frame of reference leading to a variety of salient consequences for the behavior and mental life of RIPC patients.

Thus, clinical observations complemented by the methods of neurovisualzation and spectral DCM opened the way to the current progress in understanding of self-referential cognition and, potentially, its integration with emotional experience. Indeed, little attention was devoted in the discussion to the emotional life of our patients. Our data did not support the hypothesis about division of labor between hemispheres based on the emotional valence. If the mood of our LIPC patients had a reduced emotional flare this does not mean that mechanisms of positive emotions were somehow expressed by the left hemisphere but rather that these patients were able to realistic evaluation of their-own health condition. Conversely, in RIPC patients, the dominant mood was that of neutral placidity and mild euphoria, whereas no episodes of worries, fear or happiness were observed. Their deficit of self-referential and interpersonal processing explains this profile without reference to the alleged rightwards lateralization of negative emotions.

Conclusion

Patients with unilateral brain damage, either left or right, localized in posterior tertiary areas of the cortex (inferior parietal lobe and temporoparietal junction) present distinct patterns in everyday behavior, social competencies and problem solving. A systematic overestimation of task complexity even if the task was a trivial one was the main syndrome of the LIPC group. This overestimation resulted in a prolonged phase of redundant and often disruptive contemplations preceding task solution. A completely different pattern of difficulties was found in RIPC patients. Albeit their language and basic cultural skills were relatively preserved, they demonstrated serious disorders in experiencing emotions, theory of mind, metacognition and voluntary regulation of behavior. This pattern of results can be interpreted within a revisited multilevel framework (Velichkovsky, 1990; 2002). The revision concerns the fact that metacognitive functions usually ascribed to prefrontal regions are obviously related to posterior tertiary areas of right hemisphere as well.

The role of RIPC in personal appraisal can be furthermore explained by the strong asymmetry in causal connections of left and right hippocampi (Ushakov et al., 2016). Accordingly, an access to self-related data is based on the following heuristic: look for "*ego*"-related data at the center of egocentric spatial representation. Only the right hippocampus can provide such an easy-to-find gateway into what we call "subjective experience". After right-hemisphere lesions, "Self" location within the bisected spatial frame of reference may be somewhere shifted and lost preventing access to and processing of self-related information. In a sense, such an exceptional function of the right hippocampus in the self-referential processes reminds one that was once attributed to the pineal gland on the reason that it is not an anatomically duplicated part of the brain and, thus, could serve as the site of the Aristotelian *sensus communis*. In the first formulation of this theory, Descartes wrote: "And since it is the only solid part of the whole brain which is unique, it is necessary that it is the seat of the *sensus communis*, that is to say, that of thought, and as a consequence that of the soul; for the one cannot be separated from the other" (Descartes, 1640). In a similar vein, we can say that the right hippocampus is unique in its holistic representation of surrounding space, which seems to function as the common interface for the bodily "Self " as well as for the higher-order thought and feelings.

Expanding the present focus of research to interconnected brain regions, such as the amygdala and prefrontal cortex would be consistent with the view that episodic memory performance depends on a synchronization of activities in the hippocampus and its brain's environment (e.g. Fell et al., 2001). The latter includes temporal and frontal cortices with the amygdala as the major "amplifie" (McEwen, Nasca, & Gray, 2016). Recently, a priority of the right amygdala in functional connections with the ventrolateral prefrontal cortex has been reported (Kerestes, Chase, Phillips, Ladouceur, & Eickhoff, 2017). This can be a sign of more profound differences in the cause-and-effect connectivity of temporal and prefrontal cortices with the amygdalo-hippocampal region and underlying structures involved in regulation of basic needs and emotional reward of activity. In any case, the current knowledge about contrasting functions of posterior tertiary areas of left and right hemisphere will be an essential component of modeling human cognitive-affective architecture in the years to come.

Acknowledgements

This study was in part supported by the Russian Science Foundation (grants 14-28-00234 and 17-78-30029 aimed at investigation of self-referential cognition and semantic representations in the human brain, respectively) and by the Russian Foundation for Basic Research (ofi-m grants 15-29-01344 and 17-29-02518 as related to attentional control and to the large-scale network organization, respectively). We wish to thank Marie Arsalidou and Olga Efimova for their help in formulating our ideas about brain mechanisms of consciousness.

References

Arsalidou, M., Pascual-Leone, J., Johnson, J., Morris, D., & Taylor, M. J. (2013). A balancing act of the brain: Activations and deactivations driven by cognitive load. *Brain and Behavior*, *3*(3), 273–285. doi: 10.1002/brb3.128

Atchley, R.A., Story, J., & Buchanan, L. (2001). Exploring the contribution of the cerebral hemispheres to language comprehension deficits in adults with developmental language disorder. *Brain and Cognition*, *46*, 16–20. doi: 10.1006/brcg.2000.1268

Baier, B., & Karnath, H.O. (2008). Tight link between our sense of limb ownership and self-awareness of actions. *Stroke. 39*(2), 486–488. doi: 10.1161/STROKEAHA.107.495606

Balonov, L.Ja., Deglin, V.L., & Chernigovskaja, T.A. (1979). Funktionalnaja asimmetrija mozga v organizatii rechvoj aktivnosti [Functional brain asymmetry in organization of language activity]. In *Sensornie sistemy* [Sensory systems] (pp. 35–39). Leningrad: Nauka.

Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., & St. George, M. (2000). The neurocognitive profile of Williams syndrome: A complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience, 12* (Suppl. 1), 7–29. doi: 10.1162/089892900561959

Bernstein, N.A. (1990). *O postrojenii dvizhenij* [On the construction of movements]. In N.A. Bernstein, *Physiologija aktivnosti* [Physiology of activity]. Moscow: Nauka. (Original work published 1947)

Bischof-Koehler, D. (1989). *Spiegelbild und Empathie: Die Anfaenge der sozialen Kognition* [Mirror image and empathy]. Bern: Huber.

Blanke, O., & Mohr, C. (2005). Out-of-body experience, heautoscopy, and autoscopic hallucination of neurological origin: Implications for neurocognitive mechanisms of corporeal awareness and self-consciousness. *Brain Research Reviews*, *50*(1), 184–199. doi: 10.1016/j.brainresrev.2005.05.008

Buckner, R.L., Andrews-Hanna, J.R., & Schacter, D.L. (2008). The brain's default network: Anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences, 1124*, 1–38. doi: 10.1196/annals.1440.011

Burgess, C., & Lund, K. (1998). Modeling cerebral asymmetries in high-dimensional space. In M. Beeman & C. Chiarello (Eds.) *Right hemisphere language comprehension: Perspectives from cognitive neuroscience* (pp. 215–244). Mahwah, NJ: Lawrence Erlbaum.

Burgess, N., Jackson, A., Hartley, T., & O'Keefe, J. (2000). Predictions derived from modeling the hippocampal role in navigation. *Biological Cybernetics, 83*, 301–312. doi: 10.1007/s004220000172

Burgess, P.W., Cohen-Yaacovi, G., & Volle, E. (2012). Rostral prefrontal cortex. In B. Levine & F.I.M. Craik (Eds.) *Mind and the frontal lobes: Cognition, behavior, and brain imaging* (pp. 47–92). New York: Oxford University Press.

Carper, R.A., Treiber, J.M., DeJesus, S.Y., & Müller, R.-A, (2016). Reduced hemispheric aymmetry of white matter microstructure in Autism Spectrum Disorder. Journal of the American *Academy of Child & Adolescent Psychiatry, 55*(12), 1073–1080. doi: 10.1016/j.jaac.2016.09.491

Chiarello, C. (1998). On code of meaning and the meaning of codes: Semantic access and retrieval within and between hemispheres. In M. Beeman & C. Chiarello (Eds.) *Right hemisphere language comprehension: Perspectives from cognitive neuroscience* (pp. 141–160). Mahwah, NJ: Lawrence Erlbaum.

Cohen, M.S., Kosslyn, S.M., Breiter, H.C., DiGirolamo, G.J., Thompson, W.L. Anderson, A.K., … Belliveau, J.W. (1996). Changes in cortical activity during mental rotation: A mapping study using functional magnetic resonance study. *Brain*, *119*, 89–100. doi: 10.1093/brain/119.1.89

Craik, F.I.M., Moroz, T.M., Moscovitch, M., Stuss, D.T., Winocur, G., Tulving, E., & Kapur, S. (1999). In search of the Self: A positron emission tomography study. *Psychological Science*, *10*(1), 26–34. doi: 10.1111/1467-9280.00102

Descartes, R. (1640). Letter to Mersenne, 24 December 1640. In: C. Adam & P. Tannery (Eds.) (1964–1974). *Oeuvres de Descartes* (vol. III, no. 223), 2nd ed., Paris: Vrin.

Dickerson, B. C., & Eichenbaum, H. (2010). The episodic memory system: Neurocircuitry and disorders. *Neuropsychopharmacology Reviews*, *35*, 86–104. doi: 10.1038/npp.2009.126

Dolina, I.A., Efimova, O.I., Kildyushov, E.M., Sokolov, A.S., Khaitovich, P.E., Nedoluzhko, A.V., … Velichkovsky, B.M. (2017). Exploring *terra incognita* of cognitive science: Differential gene expression in the most rostral parts of human brain. *Psychology in Russia: State of the Art, 10*(3), 231-247. doi: 10.11621/pir.2017.0316

El-Gaby M., Shipton, O.A., & Paulsen, O. (2015). Synaptic plasticity and memory: New insights from hippocampal left-right asymmetries. *Neuroscientist*, *21*(5), 490–502. doi: 10.1177/1073858414550658.

Forgács, B., Bohrn, I., Baudewig, J., Hofmann, M.J., Pléh, C., & Jacobs. A.M. (2012). Neural correlates of combinatorial semantic processing of literal and figurative noun compound words. *NeuroImage*, *63*, 1432–1442. doi: 10.1016/j.neuroimage.2012.07.029

Forgács, B., Lukács, Á., & Pléh, C. (2014). Lateralized processing of novel metaphors: Disentangling figurativeness and novelty. *Neuropsychologia*, *56*, 101–109. doi: 10.1016/j.neuropsychologia.2014.01.003

Federmeier, K.D., Mai, H., & Kutas, M. (2005). Both sides get the point: Bihemispheric sensitivity to sentential constraint. *Memory and Cognition*, *33*, 871–886. doi: 10.3758/BF03193082

Gazzaniga, M.S., (Ed.). (2009). *The cognitive neuroscience*. 4th Ed. Boston, MA: MIT Press.

Gelb, A., & Goldstein, K. (1920). *Psychologische Analysen hirnpathologischer Faelle* [Psychological analyses of brainpathology cases]. Leipzig: Barth.

Gusnard, D.A., Akbudak, E., Shulman, G.L., & Raichle, M.E. (2001). Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *PNAS, 98*, 4259–4264. doi: 10.1073/pnas.071043098

Gusnard, D.A., & Raichle, M.E. (2001). Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience, 2*, 685–694. doi: 10.1038/35094500.

Head, H. (1920). *Studies in neurology*. Oxford: Oxford University Press.

Heilman, K.M. (2014). Possible mechanisms of anosognosia of hemiplegia. *Cortex*, *61*, 30–42. doi: 10.1016/j.cortex.2014.06.007

Howard, I.P., & Templeton, W.B. (1966). *Human spatial orientation*. London: Wiley.

Karenina, K., Giljov, A., Ingram, J., Rowntree, V. J. & Malashichev, Y. (2017). Lateralization of mother–infant interactions in a diverse range of mammal species. *Nature Ecology & Evolution*, *1*, 0030. doi: 10.1038/s41559-016-0030

Kerestes, R., Chase, H.W., Phillips, M.L., Ladouceur, C.D., & Eickhoff, S.B. (2017). Multimodal evaluation of the amygdala's functional connectivity. *NeuroImage, 148*, 219–229. doi: 10.1016/j.neuroimage.2016.12.023

Khomskaja, E.D. (1987). *Neuropsikhologia* [Neuropsychology]. Moscow: Prosveshchenie.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T.C. & Friederici, A.D. (2004). Music, language, and meaning: Brain signatures of semantic processing. *Nature Neuroscience, 7*, 302–307. doi: 10.1038/nn1197

Koffka, K. (1935). *Principles of Gestalt psychology*. New York: Harcourt, Brace & World.

Koivisto, M., & Laine, M. (2000). Hemispheric asymmetries in activation and integration of categorical information. *Laterality*, *5*, 1–21. doi: 10.1080/713754358

Kolb, B., & Whishaw, I.Q. (2015). *Fundamentals of human neuropsychology*. 7th Ed. New York: Worth Publishers.

Korsakova, N.K., & Moskovichute, L.I. (1988). *Klinicheskaja neuropsihologija* [Clinical neuropsychology]. Moscow: Moscow University Press.

Kortte, K.B. Wolfman McWhorter, J., Pawlak, M.A., Slentz, J., Sur, S., & Hillis, A.E. (2015). Anosognosia for hemiplegia: The contributory role of right inferior frontal gyrus. *Neuropsychology, 29*(3), 421–432. doi: 10.1037/neu0000135

Krotkova, O.A., & Velichkovsky, B.M. (2008). Mezhpolusharnije razlichija myshlenija pri porazhenijah visshih gnosticheskih otdelov mozga [Hemispheric differences of thinking in the damages of higher gnostic brain regions]. In B.M. Velichkovsky & V.D. Soloviev (Eds.). *Komputery, mozg, poznanie* [Computers, brain, cognition] (pp. 107–132). Moscow: Nauka.

Luria, A.R. (1966). *Higher cortical function in man*. NewYork: Basic Books.

Luria, A.R., & Tsvetkova, L.S. (1966). *Neuropsychologicheskij analis reshenija zadach* [Neuropsychological analysis of problem solving]. Moscow: Moscow University Press.

McEwen, B.S., Nasca, C., & Gray, J.D. (2016). Stress effects on neuronal structure: Hippocampus, amygdala, and prefrontal cortex. *Neuropsychopharmacology Reviews, 41*(1), 3–23. doi: 10.1038/npp.2015.171

Mehta, Z., & Newcombe, F. (1991). A role for the left hemisphere in spatial processing. *Cortex*, *27*(2), 153–167. doi: 10.1016/S0010-9452(13)80121-9

Moser, E.I., & Moser, M.B. (2008). A metric for space. *Hippocampus, 18*, 1142–1156. doi:10.1002/hipo.20483.

Ocklenburg, S., & Güntürkün, O. (2012). Hemispheric asymmetries: The comparative view. *Frontiers in Psychology, 3*, 1–9. doi: 10.3389/fpsyg.2012.00005

Penfield, W., & Evans, J. (1935). The frontal lobe in man: A clinical study of maximum removals. *Brain*, *58*, 115–133. 10.1093/brain/58.1.115

Radlov, N.E., Harms, D., Dilatorskaja, N.L., & Gernet, N. (2015). *Rasskazy v kartinkakh* [Stories in pictures]. Moskva: Melik-Pashaev Publishing House. (Original work published 1937)

Raichle, M.E., MacLeod, A.M., Snyder,A.Z., Powers,W.J., Gusnard,D.A., & Shulman, G.L. (2001). A default mode of brain function. *PNAS, 98*, 676–682. doi: 10.1073/pnas.98.2.676

Robinson, J.L., Salibi, N., & Deshpande, G. (2016). Functional connectivity of the left and right hippocampi: Evidence for functional lateralization along the long-axis using meta-analytic approaches and ultra-high field functional neuroimaging. *Neuroimage, 27*, 64–78. doi: 10.1016/j.neuroimage.2016.04.022

Schilbach, L., Eickhoff, S.B., Rotarska-Jagiela, A., Fink, G.R., & Vogeley, K. (2008). Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the "default system" of the brain. *Conscious Cognition, 17*, 457–67. doi: 10.1016/j.concog.2008.03.013

Shammi, P., & Stuss, D.T. (1999). Humor appreciation:A role of the right frontal lobe. *Brain*, *122*, 657–666. doi: 10.1093/brain/122.4.657

Sharaev ,M., Zavyalova, V., Ushakov, V.L., Kartashov, S.I., & Velichkovsky, B.M. (2016). Effective connectivity within the default mode network: Dynamic causal modeling of resting-state fMRI data. *Frontiers in Human Neurosciences*, *10*, 14. doi: 10.3389/fnhum.2016.00014

Shipton, O.A., El-Gaby, M., Apergis-Schoute, J., Deisseroth, K., Bannerman, D.M., Paulsen, O., & Kohl, M.M. (2014). Left-right dissociation of hippocampal memory processes in mice. *PNAS*, *111*(42), 15238–41523. doi: 10.1073/pnas.1405648111.

Singh-Curry, V., & Husain, M. (2009). The functional role of the inferior parietal lobe in the dorsal and ventral stream dichotomy. *Neuropsychologia*, *47*, 1434–1448. doi: 10.1016/j.neuropsychologia.2008.11.033

Sokolov, E.N. (2013). *Psychophysiology of consciousness*. New York: Oxford University Press.

Stuss, D.T., Rosenbaum, R.S., Malcom, S., Christiana, W. & Keenan, J.P. (2005). The frontal lobes and self-awareness. In T.E. Feinberg & J.P. Keenan (Eds.), *The lost Self: Pathologies of the brain and identity* (pp. 50–64). NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195173413.003.0005

Ushakov, V.L. Sharaev, M.G., Kartashov, S.I., Zavyalova, V.V., Verkhlyutov V.M., & Velichkovsky, B.M. (2016). Dynamic Causal Modeling of hippocampal links within the human default mode network: Lateralization and computational stability of effective connections. *Frontiers in Human Neuroscience*, *10*, e64466. doi: 10.3389/fnhum.2016.00528

Vallar, G., Bottini, G. & Sterzi, R. (2003). Anosognosia for left-sided motor and sensory deficits, motor neglect, and sensory hemiinattention: Is there a relationship? *Progress in Brain Research*, *142*, 289–301. doi: 10.1016/S0079-6123(03)42020-7

Velichkovsky, B.M. (1990). The vertical dimension of mental functioning. *Psychological Research, 52*, 282–289. doi: 10.1007/BF00877536

Velichkovsky, B.M. (2002). Heterarchy of cognition: The depths and the highs of a framework for memory research. *Memory, 10*(5/6), 405–419. doi: 10.1080/09658210244000234

Velichkovsky, B.M. (2007). Towards an evolutionary framework for human cognitive neuroscience. *Theoretical Biology*, *2*(1), 3–8. doi: 10.1162/biot.2007.2.1.3

Velichkovsky, B.M., Joos, M., Helmert, J.R., & Pannasch, S. (2005). Two visual systems and their eye movements: Evidence from static and dynamic scene perception. In B. G. Bara, L. Barsalou & M. Bucciarelli (Eds.), *Proceedings of the XXVII Conference of the Cognitive Science Society* (pp. 2283–2288). Mahwah, NJ: Lawrence Erlbaum.

Velichkovsky, B.M., Klemm, T., Dettmar, P., & Volke, H.-J. (1996). Evozierte Kohaerenz des EEG II: Kommunikation der Hirnareale und Verarbeitungtiefe [Evoked coherence of EEG II: Communication of brain areas and depth of processing]. Zeitschrift fuer EEG-EMG, 27, 111–119.

Vincent, D.J., Bloomer, C.J., Hinson, V.K., & Bergmann, K.J. (2006). The range of motor activation in the normal human cortex using bold FMRI. *Brain Topography, 18*, 273–280. doi: 10.1007/s10548-006-0005-y

Wilkins, A., & Moscovich, M., (1978). Selective impairment of semantic memory after temporal lobectomy. *Neuropsychologia*, *16*, 73–79. doi: 10.1016/0028-3932(78)90044-1

Winner, E., & Gardner, H., (1977). The comprehension of metaphor in brain-damaged patients. *Brain*, *100*, 717–729. doi: 10.1093/brain/100.4.717

Original manuscript received August 25, 2017
Revised manuscript accepted September 11, 2017
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society | Lomonosov Moscow State University

On the challenge of polysemy in contemporary cognitive research: What is conscious and what is unconscious

Vera I. Zabotkina[a*], Elena L. Boyarskaya[b]

[a] *Russian State University for the Humanities, Moscow, Russia*

[b] *Immanuel Kant Baltic Federal University, Kaliningrad, Russia*

* Corresponding author. E-mail: zabotkina@rggu.ru

Background. The problem of polysemy has attracted scholars' attention since antiquity and interest in the phenomenon never lessens. A substantial number of works have been published on the cognitive nature of meaning ambiguity. Despite a new emphasis on the cognitive aspects of polysemy, little has been done towards an integrated approach to the study of this linguistic phenomenon.

Objective and Method. This work's objective was to contribute to an integrated interdisciplinary theory of polysemy. To this end, we explored the cognitive foundation of meaning using empirical and theoretical research methods, but mostly relying on semiotic analysis of texts central to the humanities. In particular, we analyzed the dichotomy of conscious vs. unconscious processing in the acquisition and use of polysemy. For the identification of cognitive patterns of polysemy development in ontogenesis, we used probabilistic conceptual modeling.

Results. The acquisition of meaning is a conscious process: it is a conscious interaction of the speaker with an interlocutor and their common social environment. On the other hand, meanings are unconscious unless a connection between the phonological, acoustic form and the concept is established. Correspondingly, polysemy is conscious when a new meaning is formed in the course of social interaction. However, polysemy, as an inherent language phenomenon, remains unconscious for native speakers, who are unaware of its presence provided they are not involved in some form of intentional language games (pun, zeugma or intended ambiguity).

Conclusion. The present approach to the analysis of meaning ambiguity seems to be a productive endeavor. Further research into polysemy has to be based on a range of additional types of evidence, including those obtained by methods of cognitive neuroscience.

Keywords: cognition, polysemy, meaning acquisition, development, language games, concept, consciousness, unconscious processes

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0302
http://psychologyinrussia.com

Introduction

The problem of polysemy has attracted scholars' attention since antiquity, and the interest in this phenomenon, "the wild world of polysemy … its apparent semantic chaos" (Pinker, 2007, pp. 112-113), never lessens. Despite a new emphasis on the cognitive aspects of polysemy, little in the way of an integrated approach to the study of this language phenomenon has been done. This work intends to contribute to such an integrated theory. The very term *polysemy* suggests that one may achieve a much better understanding of what the meaning of a word really is, by the search for answers to the following questions: a) what is the primary meaning of a polysemous word and how is it acquired? b) is there a difference in the mechanisms of acquisition of the primary and secondary word meanings? c) is there a difference in the mechanisms of acquisition of word meanings in adults and children? d) what aspects of polysemy may be conscious, and what unconscious?

Questions about the nature of meaning present fundamental challenges, not only for linguistics, but also, when integrated into an interdisciplinary research paradigm, for philosophy, psychology, and artificial intelligence. There are various types of integrated knowledge that emerge as a result of interaction among the different sciences constituting cognitive science: integrated methodological knowledge, integrated empirical knowledge, and integrated theoretical knowledge. Cognitive semantics is the epitome of this integrative approach, the integration of integrations (Zabotkina, 2016).

It is cognitive semantics, often referred to as conceptual semantics, that allows researchers to break the gridlock in three trends in linguistic research — linguistic determinism, nativism, and radical pragmatism. Linguistic determinism assumes that language and its structure predetermine the nature and character of basic cognitive processes — categorization, perception, etc. — and impose certain limits on the learning process as such. Hence it follows that representatives of different cultures think differently. L. Wittgenstein believed that "the limits of my language mean the limits of my world" (Wittgenstein, 2016). Proponents of nativism think that the human conceptual system, mental processes, and structures are congenital in nature. Radical pragmatics, on the contrary, postulates that the word can mean almost anything depending on the context in which it is used. There is a certain rationale in each of the three schools of thought, but each of them contradicts the other two. The difference between languages — the main argument used by determinists — does not fit into the framework of nativism. Neither does polysemy, the main object of research for radical pragmatists, fit into the mold of determinism. Only cognitive semantics provides a way out of this impasse.

According to the latest research, the semantic system is organized into intricate patterns that seem to be consistent across individuals. Most areas within the semantic system represent information about specific semantic domains, or groups of related concepts. Researchers used a new generative model to create a detailed semantic atlas, the Brain Dictionary, showing which domains are represented in each area of the brain (Huth, de Heer, Griffiths, Theunissen, & Gallant, 2016). This study convincingly demonstrates that data-driven methods provide a powerful and efficient means for mapping functional representations in the brain.

However, cognitive linguistics and cognitive semantics rely on *both experimental and empirical methods* of research, which are of equal value. The skilled intuition of cognitive linguists is useful in studying specific influences of thought and embodied experience. Cognitive linguists need not become experimental psychologists or computer scientists for their work and ideas to be seen as legitimate, with considerable theoretical implications (see Gibbs, 2006, pp. 2–16). This research shows the benefits of empirical methods for the study of meaning.

Method

The objective of this research is to explore the cognitive foundation of meaning and reflect on the dichotomy of the conscious *versus* unconscious in the acquisition and use of polysemy. With this objective in mind, we refer to empirical methods of research, introspection, and probabilistic conceptual modelling in the identification of patterns and modalities of conceptual processes underlying polysemy.

We begin with a survey of previously published papers on the theory of meaning viewed from the cognitive perspective, and review several hypotheses concerning the interactions among conceptual system, thought, language, and the multiplicity of meaning. The second stage of the analysis presents empirical research into the unconscious nature of meaning, inner speech, and polysemy. We assume that the meaning of a word is based on the concept it expresses and we argue that meanings remain unconscious until a connection between the phonological, acoustic form and the concept is established. Further on, we analyze the benefits of the integrated approach to the study of polysemy using numerous examples taken from WordNet 3.1, Collins Co-Build Corpus and Dictionary, as well as contextualized samples taken from several corpora — the British National Corpus and CHILDES. The selected language material is used to identify conceptual patterns, or algorithms, of the acquisition of the primary and secondary meanings of polysemous words by children and adults.

Results

Polysemy and the challenge of meaning

Cognitive semantics requires an interdisciplinary approach to the study of meaning. One of the most significant works, in our view, is a relatively recent publication by R. Jackendoff's *A User's Guide to Thought and Meaning*, devoted to the problem of meaning addressed from the perspective of different sciences. Meanings are flexible and adaptive in nature: "The meaning of the word is the concept it expresses…. The meaning of a sentence is the thought it expresses…. Meanings are thoughts expressed by language. …They are flexible and adaptive" (Jackendoff, 2012, p. 3).

This idea dates back to the works of Lev Vygotsky, who studied the connections among concepts, words, sense, and meaning: "We found the unit that reflects the unity of thinking and speech in the meaning of the word…. That is, we cannot say that word meaning is a phenomenon of either speech or thinking. The word without meaning is not a word, but an empty sound. Meaning is a necessary, constituting feature of the word itself. … *word meaning is nothing other than a generalization, that is, a concept*" (https://www.marxists.org/archive/vygotsky/works/words/ch07.htm).

In dictionaries the meanings of polysemantic words are organized in a single entry, because *lexicographers* have found something in common between the meanings, and on this basis, decided to class the word as a polysemous one. In most cases, the conceptual and consequently, semantic, connection between the meanings of polysemous words is not in doubt, but there remains a degree of subjectivity in ascribing meanings to one word. Traditional semantics considers distinguishing polysemy from homonymy as one of its fundamental areas of research. When comparing the structure of polysemous words in bilingual dictionaries, the researcher is faced with an even more difficult task of finding correlations not only between two language systems but, more importantly, between two conceptual systems that appear to be markedly different. One cannot but agree with the poet Marina Tsvetayeva: "Some thoughts are unthinkable in another language" (http://www.tsvetayeva.com/letters/let_chern).

With this knowledge, what takes place in our mental dictionary ultimately depends upon various hypotheses and assumptions. The *hypothesis of unconscious meaning* put forward by R. Jackendoff is particularly interesting. According to this researcher, word meanings fulfil two essential functions: a reference function — that meanings connect language to the world; and an inference function — that meanings serve as a vehicle of reasoning (Jackendoff, 2012, pp. 47–48).

It is common knowledge that not all concepts can be expressed in words, and not all senses can be conveyed. This is surely an argument in favor of the view that meanings are unconscious. Many languages have expressions such as "this is spinning in my head", "it is on the tip of my tongue", and "this is not what I wanted to say". The human mind clearly identifies the activation of a certain amount of conceptual information connected to a word that does not spring to mind at the right time. There is a feeling that the forgotten word is about to be recalled and will pop up in an instant. But that rarely happens. Everybody has experienced frustration over the inability to recall a word that is so familiar and so common. Consider the excerpt from O. Mandelstam's poem "The Swallow", which poetically describes the situation: "I have forgotten the word I meant to say/and the *voiceless thought*/ returns to the castle of shadows" (http://www.poetryintranslation.com/PITBR/Russian/Mandelstam).

In these cases, the brain fails to establish a connection between the acoustic form of the word and the concept it expresses. The situation described above often has a continuation — the lost word is recalled much later, as if out of "nowhere", when people are already engaged in another cognitive activity. The brain continues to work on resolving the problem *without us being conscious* of the task, without us being *aware* of it. The brain solves the task working in "background" mode.

There is a significant number of concepts that have no names attributed to them. According to S. Pinker, there are concepts that simply refuse to have names: "a concept that everyone wants to express, but for which *le mot juste* does not yet exist. … Many gaps in the language simply refused to be filled … the lout sitting next to you on a train or in an airport lounge who screams into a cell phone the whole time. The disgusting lumps of brown snow that accumulate behind a car wheel… etc." (Pinker, 2008, pp. 304–305). There are no names, aka words, for some fragments of experience, emotions, feelings, or situations, which are familiar to everyone, and strange as it may seem, for objects having no names. Even if the concept has a name, the entire amount of conceptual information associated with the word is not fully reflected in its meaning.

The acoustic form of a word is activated in three cases: during its acoustic perception, during speech production, or in inner speech. Inner speech has a well-defined form and is observed during different phases of brain activity. Deaf speakers perceive inner speech as a sign language image. According to Jackendoff's hypothesis: "Pronunciation is conscious … and it is linked to unconscious meaning — the thought or concept that the pronunciation expresses" (Jackendoff, 2012, p. 86). Pronunciation results from a physical, acoustic process: to pronounce means "to say, speak or utter something in a certain way" (WordNet). It is impossible to say a word and not to be aware of it, at least of its acoustic form. Sounds belong to the physical world, i.e., the external environment. However, the "acoustic form" also exists in inner speech.

Inner speech is certainly a remarkable phenomenon. L. Vygotsky proposed a theory of inner speech describing its main characteristics. He thought that "in inner speech, we find a predominance of the word's sense over its meaning. A word's sense is the aggregate of all the psychological facts that arise in our consciousness as a result of the word. Sense is a dynamic, fluid, and complex formation which has several zones that vary in their stability. Meaning is only one of these zones, that of the sense that the word acquires in the context of speech" (Vygotsky, 1934, pp. 248–249).

Vygotsky's ideas are of cardinal importance for the cognitive study of polysemy. Of interest is his observation of the dynamics of sense and meaning in different contexts. Vygotsky described inner speech as a complex, dynamic process involving transformation of its predicative, idiomatic structure into syntactically articulated speech intelligible to others. In inner speech, thoughts are verbalized and verbal thoughts take the following course: from the motive that engenders a thought to the shaping of the thought, first in inner speech, then in meanings of words, and finally in words (ibid).

However, it is not clear how to interpret the inner speech that people "utter and hear" in their dreams. Dreams are "whispers of the unconscious", so how can the brain be conscious of pronunciation and the general flow of communication? In his "Psychological Notes", V. Odoevsky, a Russian writer of the 19th Century, describes an amazing creature that he saw in a dream — "a compound of darkness, death, and a minor chord. It is impossible to express it verbally after waking up, but in the dream the creature had a name and the name was clear to me" (Odoevsky, 1975). Will such an acoustic form of the name be considered as *perceived, and thus, conscious*? Are word meanings still *unconscious* in the unconscious?

The conceptual scope and meanings of polysemous words are *unconscious* for native speakers. They do not realize the complexity of the structure of polysemous words, the entire volume of conceptual information encoded by them. The identification of the meaning of a polysemous word occurs so effortlessly that polysemy is perceived as monosemy.

When generating an utterance, the speaker chooses words according to his/her communicative and pragmatic intention. Polysemy, as it may seem, is a problem for the listener, because he/she must adequately identify the meaning of the polysemous word in a given context. However, native speakers easily identify the meaning, and in most cases, are not aware of the presence of polysemy. Both the speaker and the listener perceive the words as *monosemous and are unconscious of the presence of polysemy.*

The only exceptions are examples of pun and zeugma, in which the speaker uses a deliberately ambiguous word in different meanings to achieve a specific pragmatic effect: "We must all hang together, or assuredly we will all hang separately" (B. Franklin), or "She opened her door and her heart to the orphan" (http://www.wunderland.com/). Our analysis has shown that in such cases, both the speaker and the listener are aware of ambiguity; they are *conscious* of the multiplicity of meanings.

Speakers are also conscious of polysemy when explaining the meanings of polysemous words, for instance, in academic discourse. The following example shows an intentional, i.e., *conscious*, cognitive effort to explain the meaning of the adjective "special":

> Q: If anyone has any ideas on how to explain the meaning of the word "special", it would be greatly appreciated.
>
> A: Do you mean "special needs", or just "special"? I'd go with "Some people need to receive extra help at school if they have physical or learning difficulties"…. dyslexia is a good example? Or if you mean just "special", it's best described as something that's better or greater. "Special" surprise = a surprise that's better than a run of the mill one. "Special" Agent = a better agent, higher ranking, etc. "Special" Fried Rice = the ultimate in fried rice.

We also argue that meanings are unconscious *until a connection between the phonological, acoustic form and the concept is established*. However, the acquisition of meaning is a *conscious process;* it is a *conscious interaction* of the speaker with another speaker and the external environment. These provisions play an important role in our study of polysemy.

We hold that the acquisition of both primary and secondary meanings of polysemous words most often involves an active cognitive modality, i.e., an interaction with another party to the communication; *it is socially and linguistically mediated* cognition. The acquisition of a word meaning is an interaction between a significant number of mechanisms — conceptual, social, and linguistic ones, each having a complex nature. Conceptual mechanisms of this conscious process are reflected in the human ability to form concepts, analyze them, and attribute the newly formed concepts to the already existing categories. Social mechanisms manifest themselves in the form of verbal and non-verbal communication, for example, when the speaker can just point to the new *signified* (the Point-and-Say method).

We have investigated situations (these samples are taken from CHILDES and other sources) when an adult speaker is explaining the meaning of a word to a child. Our research shows that this dialogue follows an almost prototypical scenario:

> Child: What is it?
>
> Adult: This is X.
>
> Child: What is X?
>
> Adult: X is **like** Z.
>
> Child: X is like Z? And can Z …….?

There are numerous examples of such dialogues in various corpora and parents fora containing samples of children's speech (CHILDES et al.). Adult speakers often use analogy, metaphor, comparison, as well as other members of the same conceptual category, helping the child form a new concept. In most cases, children continue to ask questions till they get enough information to form the concept and bind it to its acoustic form:

CHILD: *chto takoe derevo*?
/what is a tree?/
MOTHER: eto takoe mesto, na kotorom mnogo mnogo listochkov.
/this is a place where many many leaves grow/
MOTHER: I tolstyj stvol s koroj.
/And it has a trunk and bark/
CHILD: *chto takoe kra, chto takoe kor?*
/ what is ba.., what is ba..k?/
MOTHER: eto kozha u dereva.
/this is the tree's skin/
CHILD: chto takoe malen'kaja berezka?
/ what is a small birch-tree?/
(CHILDES)

In this example, the adult speaker gives the definition of the word *tree,* describing it as a "*place where many, many leaves grow*", and then provides additional information about other attributes of the concept — "*it has a trunk and bark*". This was obviously not enough for the child to form the *tree* concept. Having been asked for clarification, the adult speaker uses a metaphor, "*bark is the tree's skin*", or a simile, "bark is like skin", to better explain the meaning. The formation of the conceptual category continues, since the child asks about another member of the same category — a small birch tree.

The formation of conceptual categories and their extension is an ongoing process reflected in language and speech interactions. Consider the following example:

CHILD: a gde Mishkino *guljan'e*?
/and where is Mishka's *walkening*?/
MOTHER: na ulice
/out in the street/
CHILD: a gde Mishka *ne guljanije*?
/and where is Mishka's not *walkening*?/
CHILD: to *ni guljanije*?
/what is not *walkening*?/
MOTHER: takogo netu slova.
/there is no such word/
CHILD: chto takoe *niguljanije*?
/what is *unwalkening*?/

MOTHER: takogo slova net.
/there is no word like this/
CHILD: eto guljanije +…
/this is *walkening* +…/
CHILD: eto …
/this is…./
MOTHER: guljanije eto kogda guljajut.
/*walkening* is when somebody goes out for a walk/
CHILD: a chto takoe *nehoroshee guljan'e*?
/and what is *bad walkening*?/
(CHILDES)

In the example above, not only does the child ask for the clarification of the meaning of *guljanije*, but she also tries to find out if there is a word opposite in meaning to it — *niguljanije*. Having heard that there is no such word, the child expresses the idea of negation (*Rus*. negative particles *ne/ni*) in another way — by using the adjective *nehoroshee (Rus*. "not good; bad"). This is a *conscious* attempt to acquire the meaning and form the dichotomy good vs. bad, which is one of the first oppositions acquired by children.

So, we could argue that the formation of the concept and its attribution to the acoustic form is *conscious*. Our research has shown that there are certain patterns of cognitive operations that are typical of the acquisition of the primary meaning of a word:

- visual or acoustic input to the echoic or iconic memory;
- assessment of sufficiency/poverty of the stimulus;
- given poverty of the stimulus, initiation of a request for additional information required to overcome the poverty of the input and the stimulus;
- formation of the concept;
- binding the acoustic or visual form to the newly formed concept;
- formation of the meaning;
- storing the meaning in the mental lexicon.

Poverty of the stimulus means both the scarcity of perceptual stimuli (acoustic or visual), as well as the background of the event, the sparseness of the external environment. The following example illustrates the formation of a concept differing from the "correct" or "true" due to the poverty of the stimulus: "I recall my childhood guardian's daughter making me laugh numerous times. She thought my tortoise, **Hezakiah, was a hezakiah**. I'll bet to this day, she may call a tortoise a hezakiah. I tried to explain to her that it was his name. She understood naming her dolly, but not a reptile" (https://www.buzzfeed.com).

In this example, the poverty of the stimulus — insufficient initial input coupled with the dominance of the situational background — linked the acoustic form to the concept "tortoise". As a result of a metonymic transfer, the proper name was used as a generic noun. The acquisition of *secondary*, derived meanings depends, it seems, on the level of development of the human conceptual system and sufficient background knowledge. Consequently, language behavior in adults is different from that in children.

Indeed, adult speakers and children acquire the meanings of polysemous words differently. Adult speakers can acquire new meanings, both primary and secondary, after fewer presentations, since their conceptual system is fully established. Adult speech is dominated by verbs having a high degree of polysemy; the number of polysemous words in adult speech is much higher than that of children. Children, especially at the early stages of development, opt for nouns having fewer meanings. They want to have a name for everything in their external environment: "What is the name of the space between the bits that stick out on a comb?" (http://www.mama.com.au/14-questions-kids-ask). This is a *conscious* attempt at nomination.

Even though the following passage is an example of homonymy, it still demonstrates the tendency of children to use a separate word for each *signified*. The example below shows the urge of the child to have a separate word that can express his idea, his concept of gender differences: "We bought a water mister for our backyard. My son asked, "What do you call it if it's a girl?" (https://www.buzzfeed.com). This is an example of an erroneous categorization and attribution of human gender differences to an inanimate object ("mister — a form of address, a title" and "mister — a bottle with a nozzle for spraying a mist of water, as onto houseplants" (Collins Co-Build Dictionary). Children prefer less ambiguous words, thereby avoiding uncertainty.

Unlike adults, children reject attempts to use the same word in different meanings and prefer to use the words that are well known to them. They seek to comply with the law of the symmetry of the linguistic sign, so the same word should not be used to name different objects or artefacts.

We hold that during the acquisition of a new meaning of a polysemous word, there might be a conflict between the primary and secondary meanings: "Evie, aged seven, after a drug-awareness program at school: 'Mum, how do you smoke *a pot*? Do you stick your head in it? What are they breathing in? Won't the steam burn you if the stove is on?' " (a pot — "a deep round container used for cooking stews, soups, and other food" and pot — "sometimes used to refer to the drugs cannabis and marijuana", Collins Co-Build Dictionary) (http://www.mama.com.au/14-questions-kids-ask).

The conflict between the primary and secondary meanings of a polysemous word seems to be less apparent in adults than in children. So far we have failed to find convincing examples proving this argument.

We suggest that conceptual operations underlying secondary and further meaning acquisition develop according to the following algorithm:

- visual or acoustic input to the echoic or iconic memory;
- activation of the primary (basic) meaning;
- conflict between the basic and the new meanings;
- assessment of sufficiency/poverty of the stimulus;
- formation of a new concept;

- linking the concept to the acoustic form;
- establishing a connection between primary and secondary meanings through individual invited inferencing;
- storing the meanings in the lexicon.

Our analysis has shown that the sequence of conceptual operations presented above, accompanying the acquisition of primary and secondary meanings, is conditional upon the time and character of cognitive operations. It is still unclear how these operations occur: do they occur consecutively or do some of them develop simultaneously? The nature of these cognitive operations requires further research involving neurologists and neurolinguists.

The opposite process — identification of the meaning of polysemous words — depends on the cognitive context in which the new meaning is acquired and identified. Each polysemous word is associated with a set of dynamic cognitive contexts forming a complex multi-dimensional mental representation, which could potentially capture and store a significant amount of conceptual information, referring in fact to any number of conceptual domains that are relevant to the identification of a particular sense of the word (Zabotkina & Boyarskaya, 2012; Zabotkina & Boyarskaya 2013). Consider the following example:

"After telling my five-year-old daughter I'm excited because my favorite band is *coming to town to play*, she ran to her room and started cleaning up her toys. When I asked her what she was doing she said, 'Mommy, those are big guys and if they're *coming over to play* I do not want them stepping on my stuff. They'll break everything!' " (https://www.buzzfeed.com).

This example is revealing for several reasons: it is an illustration of a basic assumption that the primary meaning is preferred to the derived one (in this case "to play" — **"**<u>spend</u> time doing enjoyable things, such as using <u>toys</u> and taking part in games". The context of the situation serves as a prime for the activation of the basic meaning. Conceptual priming is a faster means of identifying a particular word meaning after the presentation of a prime. The results of our research show that a set of cognitive contexts (or a particular cognitive context) can act as a conceptual prime, leading to faster and more accurate identification of the target word sense. The method of probabilistic conceptual modelling of word sense disambiguation, which we suggest, clearly demonstrates the role of a particular type of cognitive context and conceptual primes in word sense disambiguation (ibid).

The mental lexicon performs an important role in polysemy resolution, since it is the mental lexicon that concentrates various types of cognitive processes connected with perception, processing, storage, retrieval, usage, and generation of knowledge. It is often understood as a system of concepts and links between them which have been formed as a result of human cognitive activity. The meanings and concepts they are based on form networks with other meanings related to them conceptually and, therefore, semantically. New meanings are not acquired in isolation. They integrate into existing conceptual networks. The more meanings are acquired, the more differentiated they are compared to other words and other meanings within the structure of the polysemous word. In the mental lexicon, a polysemous word may be represented by a complex mental representation — a set of cognitive contexts associated with different senses of the polysemous word. This mental representation may store a large volume of information belonging to different conceptual domains. This perspective is important in understanding what the meaning of a polysemous word is.

Conclusion

We have attempted to show the importance of an integrated approach to the challenge of polysemy. Such a novel approach is based on integrated knowledge emerging from the interaction of three disciplines of cognitive science — cognitive linguistics, cognitive psychology, and philosophy. It has been demonstrated that progress in the study of polysemy does not come from linguistics alone, but requires drawing on tools and methods from other cognitive paradigm disciplines. This allowed us to arrive at an interpretative hypothesis concerning the cognitive basis of meaning, analyzed within the dichotomy of the conscious *versus* the unconscious. This research has demonstrated that meanings are *unconscious till a connection between the phonological, acoustic form and the concept is established*. However, the acquisition of meaning is a *conscious process;* it is a *conscious interaction* of the speaker with another interlocutor and with the external environment. Polysemy as a multiplicity of meanings associated with one acoustic form is *unconscious* for the speaker or the listener provided they are native speakers. Polysemy is *conscious during intentional activation* of two or more meanings associated with an acoustic or visual form in puns or zeugmas.

We opted for the interaction-based approach, complementary to the brain-centered computational one, and suggested algorithms of cognitive processes of the primary and secondary meaning acquisition by children and adults. The acquisition of secondary meanings inevitably results in a conflict between the primary meaning and the derived one. Further research into polysemy should be based on a range of new types of evidence obtained by neurologists and neurolinguists.

Acknowledgements

This study was supported by the Russian Science Foundation, RScF grant 17-78-30029.

References

Casas, B., Català, N., Ferrer-i-Cancho, R. & Baixeries, J. (2014). The evolution of polysemy in child language. In E. A. Cartmill, S. Roberts, H. Lyn & H. Cornish (Eds)*, Proceedings of the 10th International Conference on Evolution of Language*. (pp. 409–410). Vienna, Austria, 14–17 April 2014. doi: 10.1142/9789814603638_0068

Gibbs, R. (2006). Why cognitive linguists should care more about empirical methods. In M. Gonzalez-Marquez, I. Mittelberg, S. Couson & M. Spivey (Eds.), *Methods in cognitive linguistics* (pp. 2–16). Amsterdam, Philadelphia: John Benjamins.

Jackendoff, R.A. (2012). *User's guide to thought and meaning.* New York: Oxford University Press.

Huth, A., de Heer, W.A., Griffiths, T.L., Theunissen, F.E. & Gallant, J.L. (2016). Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature, 532*, 453–458. doi:10.1038/nature17637

Murray, G.K., Jones, P.B., Kuh, D., & Richards, M. (2007). Infant developmental milestones and subsequent cognitive function. *Annual Review of Neurology*, *62*, 128–136. doi: 0.1002/ana.21120

Mandelshtam, O. (1920). Swallow. Retrieved from http://max.mmlc.northwestern.edu/mdenner/Demo/texts/swallow.html

Odoevskij, V.F. (1975). *Psihologicheskie zametki* [Psychological notes]. Moscow: Nauka.

Pinker, S. (2008). *The stuff of thought: Language as a window into human nature*. London: Penguin Books.

Pustejovsky, J., & Bouillion, P. (1996). *Lexical semantics in context*. Oxford: Clarendon Press.

Tsvetayeva, M. (1924). Pismo Chernovoj A.V. Retrieved from http://www.tsvetayeva.com/letters/let_chern

Vygotsky, L.S. (1934/1987). Thinking and speech. In R. W. Rieber & A. S. Carton (Eds.), *The collected works of L. S. Vygotsky* (Vol. 1, pp. 37–285). New York: Plenum.

Wittgenstein, L. (2016). *Tractatus Logico-Philosophicus*. Side-by-Side-By-Side Edition, version 0.44. Retrieved from https://people.umass.edu/klement/tlp/tlp.pdf

Zabotkina, V. (2016). Knowledge, reality, and interdisciplinary discourse. In *The 3rd International Multidisciplinary Scientific Conference on Social Sciences and Arts SGEM2016 Conference Proceedings, Book 3* (Vol.1, pp. 587–594). Albena, Bulgaria, 06-09 April 2016.

Zabotkina, V.I., & Boyarskaya, E.L. (2012). Cognitive modeling of sense disambiguation in polysemous words. *Journal of International Scientific Publications: Language, Individual and Society*, *6*, 254–265.

Zabotkina, V.I., & Boyarskaya, E.L. (2013). Sense disambiguation in polysemous words: Cognitive perspective. *Psychology in Russia: State of the Art*, *6*(3), 60–67. doi: 10.11621/pir.2013.0306

Original manuscript received August 25, 2017
Revised manuscript accepted September 11, 2017
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society · Lomonosov Moscow State University

Feeling of agency *versus* judgment of agency in passive movements with various delays from the stimulus

Ignat A. Dubynin*, Sergei L. Shishkin

National Research Center "Kurchatov Institute", Moscow, Russia

*Corresponding author. Email: ignd@mail.ru

Background. The sense of agency (SoA) provides us with the experience of being a physical agent with free will. On a phenomenological basis, SoA can be divided into sensory components (feeling of agency, FoA) and more cognitive components (judgment of agency, JoA). Both these components can be independently measured.

Objective and Method. A new method was developed to test the possibility of preserving SoA and its components in the atypical conditions of passive movements. Parameters of the participant's movement in response to a visual stimulus (reaction time, speed, and amplitude) were measured and used to control a servo that simulated the movement (executed passive movements). The scores on the psychometric scale of the agency and the event-related potentials (ERPs) were recorded for variable movement delays relative to the stimulus onset.

Results. It was found that the FoA was not present under passive movement conditions. At the same time, participants associated these movements with their own activity (JoA), even when their delay after the stimulus onset was too short to be actively reproduced. The somatosensory ERPs' amplitude decreased for the expected movements, demonstrating an inverse relationship with the agency scores. The lowest amplitude was observed when movements were actuated by another hand. The results can be explained using a predictive forward model, since the FoA was not observed in the absence of active movements. On the other hand, the ERPs' data and the presence of JoA with various delays between the stimulus and movement support the postdictive model of agency, where the leading role is assigned to prejudice and contextual knowledge related to the action.

Conclusion. It seems that the "context pressure" of the situation, demanding a mandatory response to the stimulus, forms a cognitive prediction of movements without firm sensory representation.

Keywords: action, free will, mental chronometry, passive movement, feeling of agency, judgment of agency, sense of agency, somatosensory event-related potentials (ERPs)

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0303
http://psychologyinrussia.com

Introduction

The sense of agency (SoA) is a specific inner experience that provides us with the feeling of being a physical agent with free will, and with the foundation necessary for sensing the spatial aspects of our "I" in action. It is an important part of human consciousness, forming the fundamental aspect of self-awareness (Gallagher, 2002).

There are several theoretical concepts that explain the emergence of SoA. One explanation is based on an internal feed-forward model, according to which motor commands and predictions about their sensory effect in a successful situation are stored in the memory (Blakemore, Wolpert, & Frith, 2000; Blakemore, Wolpert, & Frith, 2002; Karniel, 2002; Wolpert, Miall, & Kawato, 1998). If the predicted and actually sensed sensory effects coincide, agency is experienced. If they do not coincide, this event is perceived as external, and SoA is not experienced. This predictive feed-forward model or "comparator model" was used to explain the perceptual tuning to stimuli self-generated by subjects (Weiskrantz, Elliott, & Darlington, 1971) as a possible cause of disruption in the sense of control over their actions in schizophrenia (Blakemore et al., 2002). The theoretical idea of somatosensory cortical activity suppression due to the prediction of one's own motions and the accompanying sensory effects has been confirmed in some studies: behavioral (Weiskrantz et al., 1971; Bays, Wolpert, & Flanagan, 2005; Blakemore, Frith, & Wolpert, 1999), MEG (Hesse, Nishitani, Fink, Jousmäki, & Hari, 2010), fMRI (Blakemore, Wolpert, & Frith, 1998; Blakemore et al., 2000; Shergill, White, Joyce, Bays, Wolpert, & Frith, 2013), EEG and ERPs (Abbruzzese, Ratto, Favale, & Abbruzzese, 1981; Bernier, Burle, Vidal, Hasbroucq, & Blouin, 2009; Benazet, Thénault, Whittingstall, & Bernier, 2016; Sidarus, Vuorre, & Haggard, 2017).

Other approaches emphasize the role of prejudice and external situation signals for maintaining SoA (postdictive model). It has been shown that the priming of subjects with a prejudice relevant to the movement that is actually performed by another person leads to an evaluation of the action as one's own. For example, Wegner and Wheatley (1999) evoked a false SoA for movements that subjects did not do. The significance of environmental signals for SoA was demonstrated as well (Wegner, Sparrow, & Winerman, 2004). Later it was shown that the priming effect is significant for active voluntary movements (Gentsch & Schütz-Bosbach, 2011) and is particularly pronounced for passive involuntary movements (Moore, Wegner, & Haggard, 2009). A dependence of SoA on the subject's conviction about the existence of causal relationships between intention and an external event was also shown (Desantis, Roussel, & Waszak, 2011). This conforms well with the vast literature about the role of postdictive phenomena in various cognitive processes (for a review, see Shimojo, 2014).

It is possible to assume that both approaches are relevant to real agency mechanisms, a case made, for example, by Kumar & Srinivasan (2014, 2017), who examined the dependence of the sense of agency on the hierarchy of the management level (the upper level of the goal and the lower perceptual-motor level). Some researchers believe that in the context of SoA there is a clear distinction between the feeling of agency (FoA) and the judgment of agency (JoA) (Gallagher, 2000, 2006; Synofzik, Vosgerau, & Newen, 2008; Bayne & Pacherie, 2007). According to these authors, FoA is a low-level experience of being the agent of an action, without explicitly thinking about the action, whereas JoA relies on a conceptual, interpreting judgment about the authorship of an action, based on the notion of an apparent link between action and result. It was emphasized that these two aspects of agency may not in fact be linked. For example, an unexpected consequence of an action may not cause FoA, but both action and result may be referred to in JoA, if the prejudice and context suggest such a connection. It is believed that JoA is not associated with the sensorimotor system, but is a higher-level process of causal attribution. A good overview of the experimental paradigms and theoretical concepts of the sense of agency can be found in Moore (2016). A sense of ownership (SO) should also be noted in this context. It is a pre-reflexive feeling that “my body is moving”, regardless of whether this movement is active (voluntary) or passive (performed by another person or device) (Tsakiris, Schütz-Bosbach, & Gallagher, 2007).

Despite numerous empirical and theoretical studies, some issues require clarification. Is the sense of agency related to the physical possibility of performing a certain action? In particular, would a passive movement in response to a target stimulus be accompanied by a sense of agency if such a movement occurred within a response time that could not be actively reproduced by the subject? Is there a time delay threshold for a passive movement relative to the target stimulus that the movement cannot be perceived as an own one’ before the delay (the passive movement’s “reaction time”) exceeded that threshold? Under which conditions would the sensory brain activity related to passive movements decrease, and would such a decrease be accompanied by an enhanced sense of agency?

To answer these questions, we developed an experimental paradigm that allows us to monitor the parameters of a simple movement by the participant, and simulate it using a servo with a different, precisely defined delay. An experimental study was conducted using this new paradigm.

Method

14 healthy volunteers (10 men and 4 women) aged 18 to 38 years (24.9 ± 7.1, $M \pm SD$) took part in the experiment. All had normal or corrected-to-normal vision and were naive as to the hypothesis under investigation. All participants were introduced to the procedure and the instructions in writing and agreed to participate in the experiment. The study was consistent with the ethical standards of the Kurchatov complex of NBICS technologies and was performed according to the Helsinki Declaration (1964).

A simple motor action in response to a visual stimulus was the basic paradigm of our study. However, in all experimental conditions except the first, a physical movement was replaced by an imaginary one, and a servo-drive performed the necessary motion. The palm of the right hand was placed on the mounting platform (Fig. 1), and the index finger was fixed in a metal *holder* consisting of two halves connected by elastic material. The holder blocked all movements of the finger joints except for the metacarpophalangeal joint. Under the holder was a *copper plate* for the detection of finger lift by the contact method. The translational movement of the holder was transformed into a rotary motion by means of a flexible rod, which was affixed from one side to the bottom of the holder, and on the other side to the shaft of a potentiometer. The holder could be moved by the participant himself by raising his finger, or by a digital *servo*, whose beam transmitted the force of rotation of the servo shaft to the finger through a thread affixed to the upper distal part of the holder. 2 cm to the left of the holder, a red light emitting diode (*LED*) was placed at the platform height. Flashing of the LED was used as the target stimulus. The servo drive was covered with a white opaque case, to exclude any distraction that could be caused by its operation. Sensor data recording, LED flashing, and servo drive control were carried out by a computer (*PC*) using a special program written in the Delphi 2010 environment that communicated with the hardware through a special *driver* (Fig. 1).

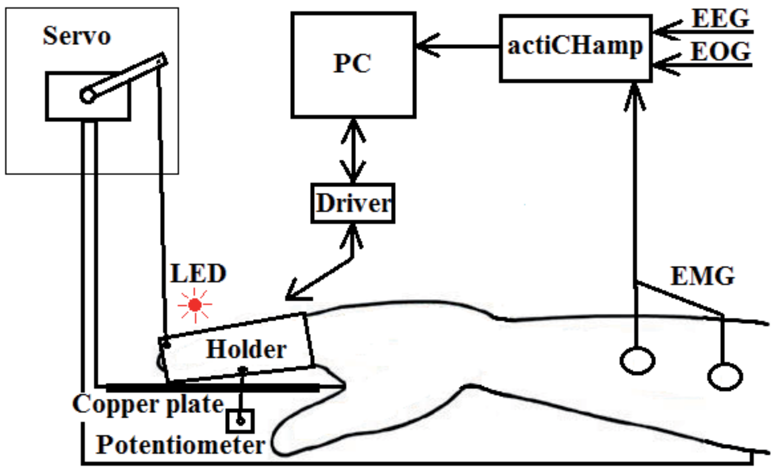


Figure 1. The experimental setup

The experiment included 6 different conditions, which were presented successively, with a 10–15 minute rest break after the 3rd condition.

Active response to a targeted stimulus without a servo (Active Movements – ActvMov). Participants were asked to carefully look at the LED and lift their finger as soon as possible when it is flashed.

Random triggering of the servo independently of the target stimulus (Stimulus Independent Passive Movements – StIndepPasMov). Participants were asked to carefully look at the LED, but instead of reacting to its flashing, they had to relax their hand as much as possible. The servo drive raised the finger randomly within the test interval, regardless of LED flashing. The distribution of servo operation events is shown in Fig. 2 (the second histogram in the upper row).

Self-raising of the finger in the holder by reacting with a free hand in response to the target stimulus (Delegated Movements – DelegatedMov). In this condition, participants had to react as quickly as possible to LED flashing by pressing the "space" key on the keyboard with their left hand, therefore initiating the servo-drive-mediated raising of the finger. The participant was asked to focus on the LED and the finger in the holder, not the one that presses the key.

Gradual decrease in the delay between the target stimulus and the operation of the servo drive (Passive Movements with Gradual Decrease – PasMovWGrdDec). The participant was asked, in response to each lighting of the LED, to imagine the finger rising movement as quickly as possible. The actual finger lifting was performed by the servo at time intervals which were gradually reduced during the experiment. For all participants, the starting delay was 230 ms from onset of the LED lighting. The delay was decreased by 5 ms every second trial.

A constant small delay between the target stimulus and the activation of the servo drive (Passive Movements with Low Delay — PasMovWLDelay). The participant had to imagine the finger movement in response to the lighting of the LED. The actual movement was carried out by the servo. The servo trip delay was 9 ± 2 ms (M ± SD).

Test with random operation of the servo drive with reference to the target stimulus (Stimulus Associated Passive Movements — StAssocPasMov). In this condition, the participant was asked to carefully look at the LED but not react to its flashing. The

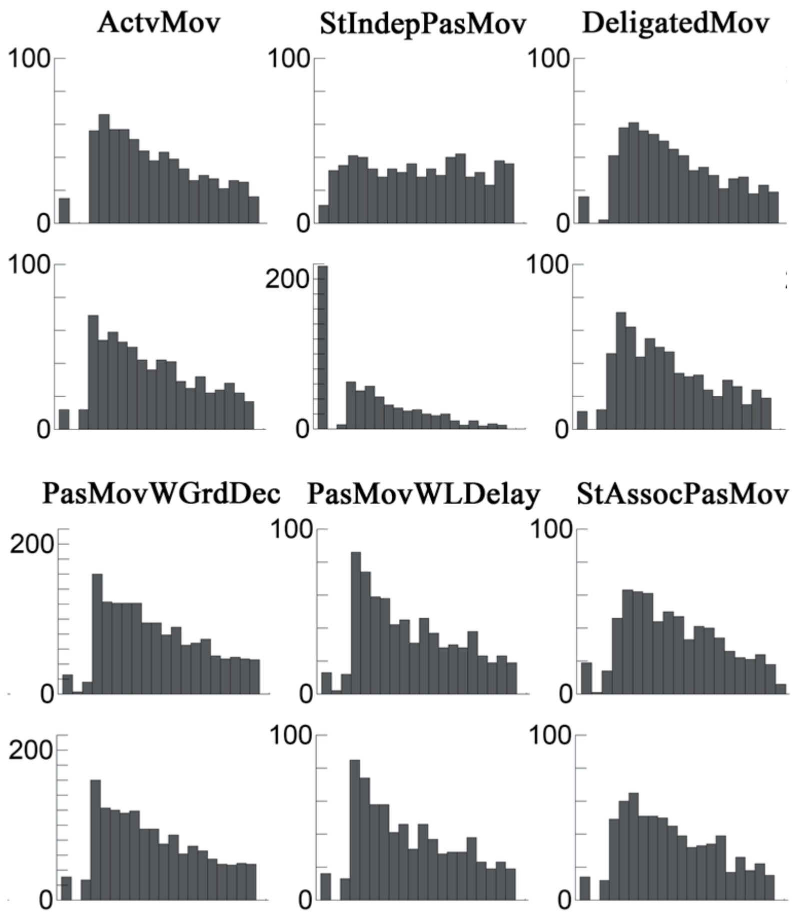


Figure 2. Histograms of the distribution of event time relative to the beginning of intervals for all six test conditions in order, from left to right. The upper row shows servo-tripping events; the bottom row shows LED flashing events. Data are collapsed over the group. X scale is –350ms … +15000 ms

servo was triggered at random intervals, but mostly after the flashing of the LED, and at times a little earlier. For 665 events, the delay time relative to the ignition of the LED was 281 ± 733 ms (min −101 ms, max 5149 ms. The distribution of the actuator trip events relative to the beginning of the intervals is shown in Fig. 2 (the rightmost histogram in the upper row).

The distribution of LED flashing under all conditions was set in such a way as to compensate for the increase in the probability of the event over time, which is inevitable when using a uniform distribution for the event generation. Stimulus time distributions for all test conditions are shown in the bottom row of Fig. 2. The probability of servo events in the *StIndepPasMov* condition was closer to a uniform distribution. A peak on the histogram for this condition was caused by servo-drive-triggering events that occurred 2 or more seconds before the planned moment of LED lighting (217 out of 648 cases, 33.5%), which led to the completion of the current interval in the absence of the stimulus. The remaining observations were distributed in the interval from −2203 to +11754 ms.

At the end of each interval, in all test conditions except for the first one (ActvMov), participants had to report verbally to the experimenter their agency score of the JoA type. A 9-point Likert-type scale (Likert, 1932) had to be used (see Table 1). The scale explanations in the table were made available to the participant on a computer screen, although, after some practice, they did not need to consult with the table in most of the trials. The experimenter noted the scores of each answer in the program.

Table 1. A psychometric scale for the judgment of agency (JoA)

Rating	Explanation
0	You did not even plan the action, there was a clear sense of an external, alien event.
1	You could perform the action, but its mental representation did not coincide at all with the actual event and was distinctly felt as not done by you.
2–3	You could perform the action, but the event actually observed largely did not coincide with its mental representation; there was a very weak sense of action ownership.
3–4	You could perform the action, but the actual observed event did not tangibly coincide with its mental representation. Some sense of action ownership.
5–6	You could perform the action, but the actually observed event slightly mismatched its mental representation; there was a pronounced sense of action ownership.
7–8	You could perform the action and the actually observed event was almost synchronous with its mental representation (a feeling of slight disagreement); you experienced a strong sense of action ownership.
9	You could perform the action and the actually observed event completely coincided with its mental representation (a feeling of complete agreement); you experienced a very strong feeling of action ownership.

Under all test conditions, the LED was flashed randomly within an interval of 15 sec duration (one trial) for a time equal to 100 ms. The beginning of each interval was preceded by a short “tick” warning sound. The sound was presented from two speakers on the table in front of the subject at a distance of about a meter, symmetrically on the left and right sides. 200 ms after the start of either active or passive movement, a confirmation sound was heard, in the form of a consonant or dissonant chord. In the test condition with active motion, the consonant chord sounded if subjects could lift their finger within 350 ms from the stimulus onset; otherwise, the dissonant chord was presented. In all other test conditions, the consonance chord was always used as the confirmation sound.

In the conditions requiring verbal reports from the subject, 2 sec after the start of the movement there was a short pause, when the experimenter recorded the score for the trial. Each condition lasted 9 min, except for the test with a gradual decrease in the delay between the target stimulus and the servo drive (*PasMovWGrdDec*), which lasted 15 min.

The logic underlying the order of the experimental conditions was as follows. The condition for a "physical" response to a target stimulus without a servo drive (*ActvMov*) was intended for getting used to the experimental conditions, and for measuring the speed and amplitude of the finger response motion. Mean values of amplitude and speed obtained during this test were used as servo parameters throughout the rest of the experiment.

StIndepPasMov and *DelegatedMov* conditions served for subjective attribution of the sense of agency to the rating scale in extreme situations, in the absence of motivation for action and unpredictability of the finger-raising events (*StIndepPasMov*). An active response to the target stimulus was combined here with imagining of the passive hand's finger movement, which was actually performed by the servo.

PasMovWGrdDec condition was designed to explore the effects of the passive movement's time delay on agency scores and ERP amplitude.

PasMovWLDelay condition was designed to test how the passive movements are perceived, with a low fixed delay between the target stimulus and the servo event.

Finally, *StAssocPasMov* was used as a control condition. The average delay of servo events in this condition was acceptable for perceiving the target stimulus and imagining the movement in response to it, but, because of the randomness of the passive movement, the participants could not accurately predict when it would start.

After the presentation of each test condition, participants were interviewed to clarify their feelings with regard to the experimental situation.

The EEG was recorded with an actiCHamp amplifier (*BrainProducts*, Germany). We used 28 electrodes with a common averaged ear reference: Fp1, F7, F3, Fz, FC5, FC1, T7, C3, Cz, CP5, CP1, P7, P3, O1, Oz, FP2, F4, FC6, FC2, FCz, T8, C4, CP6, CP2, P8, P4, Pz, O2. A vertical and horizontal electrooculogram (EOG) was recorded, as well as an electromyogram (EMG) on the right arm, where a pair of EMG electrodes was placed above the *m.extensor* at a distance of 3 cm from each other. The sampling rate was 1000 Hz. EEG and EOG were acquired in the band 0.01 Hz – 50 Hz, and EMG in the band 5 Hz – 500 Hz. The recording was made with a notch online filter 50 Hz. The electrode impedance was maintained below 10 kΩ.

The electrophysiological data were processed with the EEGLAB v13.6.5b package in the Matlab 2013b environment (*MathWorks*, USA). Statistical analysis was performed using the Statistica 10 package (*StatSoft Inc.*, USA).

Results

Behavioral results

All participants felt that there was no feeling of agency (FoA) while raising their finger in all tests except *ActvMov* and *DelegatedMov*. In the *DelegatedMov* condition, some experience of FoA of servo events appeared when the participants were told to refrain from paying attention to the hand performing the physical movement. A sense of ownership (SO), according to self-reports, was always experienced, under all test conditions.

The grand mean response time in the first test condition (*ActvMov*) was 289 ± 30 ms. The results for the JoA scale in the next two conditions were, as expected, low for *StIndepPasMov* (1.4 ± 0.8) and high for *DelegatedMov* (8.5 ± 1) (Fig. 4). In the condition with a gradual decrease in the delay between the target stimulus and the inclusion of the servo drive (*PasMovWGrdDec*), no steep changes in agency evaluations were observed, even when the servo event began to lead the LED flashes (see Fig. 3). The minimum and maximum values for group averaged agency ratings, smoothed by a moving average of 6 points, were 5.2 and 6.6, respectively, corresponding to the average delay of the passive movement start of 40 ms and 99 ms.

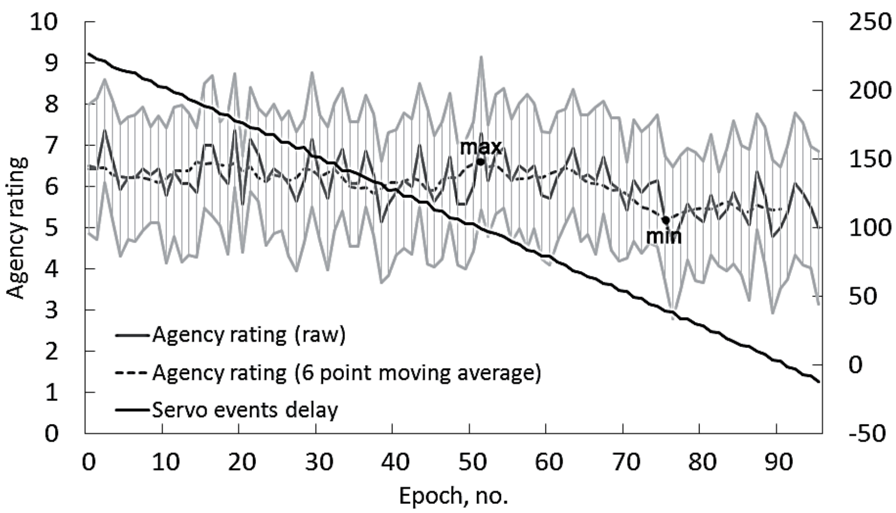


Figure 3. Grand average data of the test condition *PasMovWGrdDec*. Delays between the onset of the visual stimulus and the raising of the finger by the servo (black lines, right axis) and the agency scores (gray lines, left axis). The horizontal axis presents 96 sequential trials averaged over the group. Light gray lines show a 95% confidence interval for the agency scores.

The mean group scores and confidence intervals for four conditions are shown in Fig. 4. The effect of the condition factor was significant (Wilks' $\lambda = 0.15$, F $(3.11) = 72.97$, $p < 0.0001$). Post-hoc analysis (Fisher LSD) showed the significance of differences for all pairs of conditions ($p < 0.0001$), except for the pair *PasMovWLDelay* and *StAssocPasMov* ($p = 0.81$). The average group values of the scores of agency scale in the pair of conditions *PasMovWLDelay* and *StAssocPasMov* were very close: 5.98 ± 1.02 and 5.87 ± 0.73. These values, according to the scale of agency used in the study, point to "a pronounced sense of action ownership." The condition with a gradual decrease in the delay between the target stimulus and the activation of the servo (*PasMovWGrdDec*) was excluded from the inter-test comparative analysis, as its parameters differed too much from the other conditions.

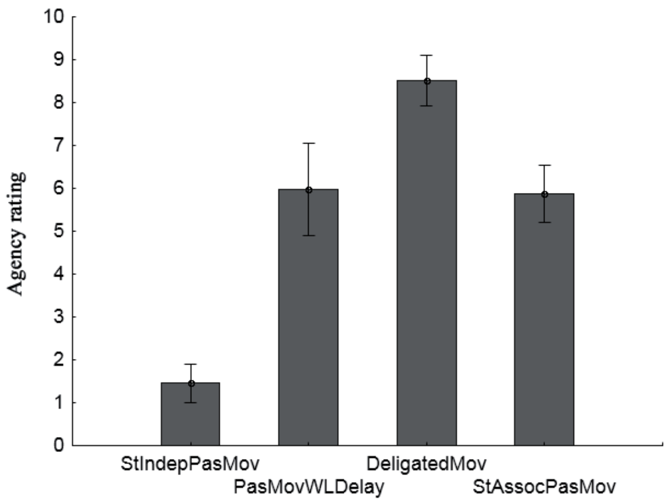


Figure 4. Average scores of the agency scale for the group (vertical scale) for test conditions (horizontal scale). Vertical lines denote 95% confidence intervals.

Electrophysiological results

The EEG and EOG data were preprocessed in the following order: filtering with FIR filter (filter order 6601) in the band 0.5 Hz – 20 Hz; extraction of epochs; baseline correction (–1000 ms to –500 ms); removal of artifact epochs using the EEGLAB function *pop_autorej* with default parameters (in particular, rejection of all epochs with an amplitude exceeding ± 100 μV) with subsequent visual inspection and additional manual rejection. For the EMG, only baseline correction and artifacts rejection were performed.

In most of the conditions, a prominent ERP negative component was observed. The maximum of this component was located at FC1 (contralateral to the involved limb), where its latency was about 115 ms. The ERP topographical scalp maps for different conditions are shown in Fig. 5 (the number of individual averages in each condition was about 40-45 epochs per participant). Based on its latency and topography, this component was identified as the somatosensory N1.

For further analysis, we chose the region of interest (ROI), for which the maximum severity of the negative somatosensory component of ERPs was observed (the ROI is marked with an oval on the map at the far right in Fig. 5). In the group analysis of the ROI, averaged data for all test conditions at mean inter-peak amplitudes were calculated as the difference between the peak amplitude of the negative ERP components (latency about 115 ms) and positive ERP components (latency about 200 ms), measured as the minimum in the 50–200 ms interval and the maximum in the 100–300 ms interval, respectively.

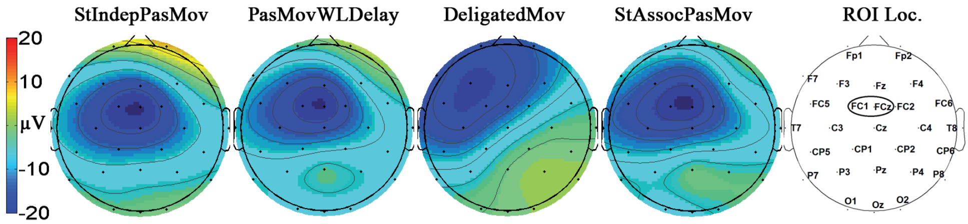


Figure 5. Topographical scalp maps of the grand average ERPs at a time point corresponding to the maximum of the somatosensory component (115 ms). Electrode positions used in the experiment are shown on a scalp map on the right. The oval on the map marks the region of interest (ROI).

In the test condition *PasMovWGrdDec*, the peak-to-peak amplitudes were calculated separately for each of the 96 EEG epochs corresponding to the monotonically decreasing servo event delays, group-averaged and smoothed with a 6 point moving average (Fig. 6). The amplitude showed a monotonic increase over most of the delay values, with a maximum delay (59 µV) at 49 ms and a minimum delay (41 µV) at 208 ms.

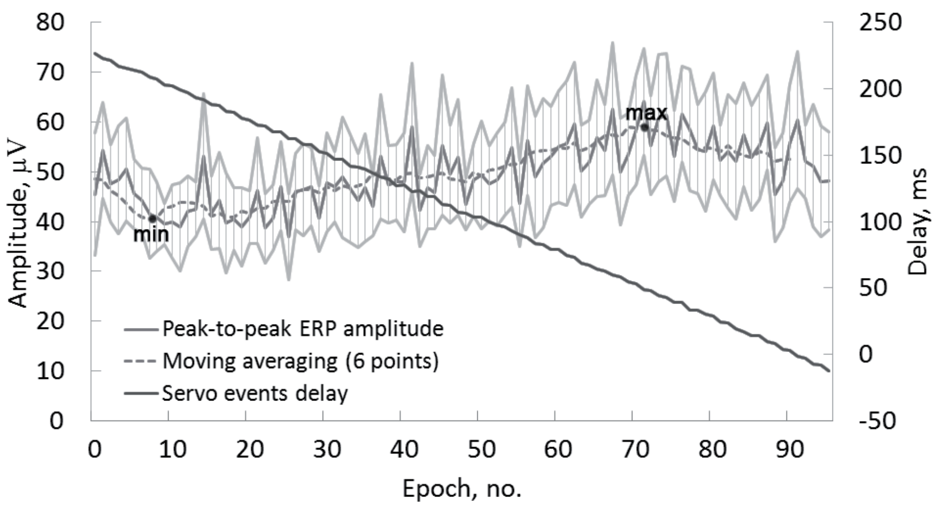


Figure 6. Group average data of the test condition *PasMovWGrdDec*. Delays between the visual stimulus onset and the raising of the finger by the servo (black lines, right axis) and the peak-to-peak amplitudes of ERP by ROI (gray lines, left axis). The horizontal axis presents 96 sequential trials averaged over the group. Light gray lines show a 95% confidence interval for the peak-to-peak amplitudes.

ROI averaged group evoked responses for the remaining five test conditions are presented in Fig. 7A. The ERP waveforms in test *ActvMov* condition was very different from the others, which was expected due to the fundamental difference in this condition — with physical movement, as against imaginary movement in the other cases. Fig. 7B shows the averaged EMG responses for the same test conditions. Note that the EMG activity amplitude for the *ActvMov* condition was noticeably different from the others.

EMG amplitude was much lower (almost absent) in all the other conditions, confirming that participants followed instructions and refrained from actively making movements when all that was required was to imagine an active role in making the movement.

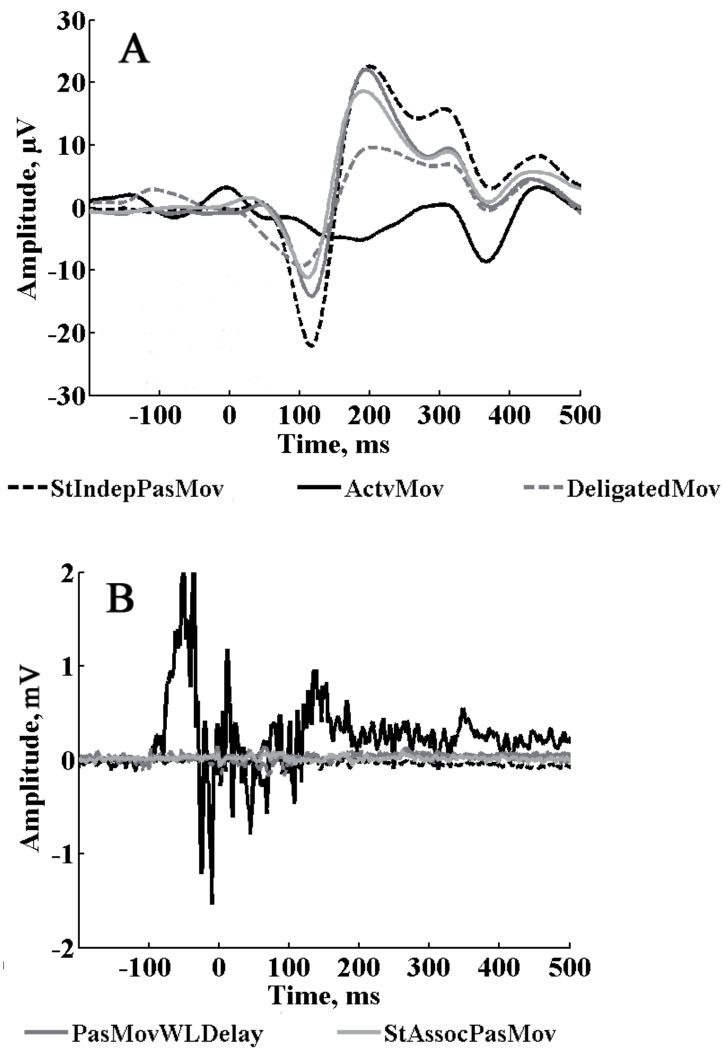


Figure 7. A: Group average ERP for ROI for different test conditions;
B: Group average EMG potentials

The mean values and the confidence intervals of the ERP somatosensory component peak-to-peak amplitude in different conditions are shown in Fig. 8.

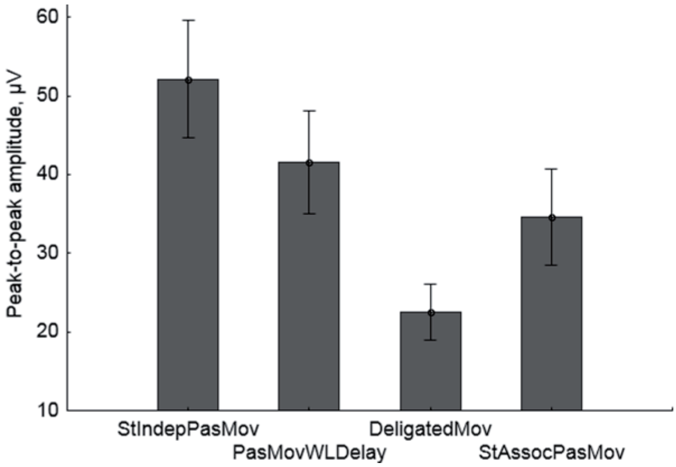


Figure 8. Grand average values of the peak-to-peak amplitude of the ERP somatosensory component (vertical scale) for passive movement conditions. Vertical lines denote 95% confidence intervals.

ANOVA with repeated measures applied to the peak-to-peak amplitude of the ERP somatosensory component showed significant effect of the condition factor (Wilks' $\lambda = 0.12$, $F(3,11) = 26.40$, $p < 0.0001$). According to the post-hoc test (Fisher LSD), all pairwise differences between the conditions were statistically significant ($p = 0.013$ for the pair *PasMovWLDelay* — *StAssocPasMov* and $p < 0.0001$ for the other five pairs).

In the passive movement conditions with imagining an active role in making the movement, the group average data for agency scores (Fig. 4) and peak-to-peak ERP amplitudes (Fig. 8) appeared to be inversely correlated: the higher the agency score in a condition, the less the peak-to-peak ERP. An ANCOVA applied to the peak-to-peak ERP amplitude as the dependent variable, with the conditions as a categorical predictor and the agency scores as a continuous predictor, showed only the influence of the condition factor ($F(3,51) = 4.53$, $p < 0.0069$), while the agency score factor effect was not significant ($p < 0.6798$).

In group averaged and smoothed-over trials, peak-to-peak ERP amplitudes (gray graph in Fig. 6) and agent scores (gray graph in Fig. 3) are slightly correlated within the *PasMovWGrdDec* condition (Spearman's $R = -0.29$; $p = 0.003$). This could reflect the influence of the common factor of the passive movement delay, which correlated negatively with the ERP amplitude (Spearman's $R = -0.68$, $p < 0.001$) and positively with the agency scores (Spearman's $R = 0.49$, $p < 0.001$).

Discussion

We investigated SoA in a group of healthy participants, asking them to imagine that they were actively moving their finger, although actually it was moved by a mechanical device, with a different time delay relative to a simple visual stimulus. One of our assumptions was that under these conditions the participants would associate the passive movements with their own intention, perceiving the actual movement as actively made by them, even if it was performed earlier than their earliest possible reaction time. Contrary to this assumption and in accordance with the survey results, in all test conditions with passive movement there was no apparent FoA. However, the participants' responses corresponded to a perception that these actions could have been performed by them; moreover, they typically chose scores related to the statement that the action only "slightly mismatched its mental representation" (except for the *StIndepPasMov* condition, in which there was no sense of agency). The sense of ownership (SO) was experienced in all test conditions. This confirms the idea that the SoA for any movement includes the SO as its most basic aspect (Gallagher, 2000), which can be also accompanied by two further aspects, FoA and JoA (Synofzik, Vosgerau, & Newen, 2008; Bayne & Pacherie, 2007).

In the *PasMovWGrdDec* and *PasMovWLDelay* conditions, the scores did not reduce to the minimum values even when it was physically impossible to perform the action that was executed by a servo, because the delay between the target stimulus and the passive movement was too small. Interestingly, no abrupt changes were observed in the group average of agency scores or in their individual dynamics that might be related to a threshold time given from the stimulus such that the action cannot be perceived as fully "owned" if it appears earlier than the threshold time. Assuming that the scores of the scale we used really reflected the JoA phenomenon, it turns out that this aspect of SoA persists with any delay between the target stimulus and the action.

The somatosensory ERP complex observed in this study apparently did not differ from a typical response to tactile stimulation (see, for example, Eimer, Maravita, Van Velzen, Husain, & Driver, 2002). It was pronounced under passive movement conditions and absent in the active movement condition. In all conditions when the passive movement was expected, the amplitude of the ERP somatosensory component decreased, possibly as the result of the suppression of somatosensory brain activity. JoA was inversely related to the magnitude of this activity. For example, with a gradual decrease in the delay in the *PasMovWGrdDec* condition, agency scores decreased along with an increase in somatosensory ERP complex peak-to-peak amplitude. Thus, active movement, executed with muscle activity, seemed to be not so important for JoA, as compared to compliance with expectations: this is what really mattered.

The results obtained on the whole allow us to conclude that peak-to-peak amplitude of somatosensory ERP reflects the unexpectedness of the passive movement event quite well. At the same time, it turned out that the scores on the scale of agency effectively differentiate the extreme conditions for predictability of events: with self-generation of the servo event by the free hand (*DelegatedMov*) and in the condition of random triggering of the servo, regardless of the target stimulus (*StIndepPasMov*). Meanwhile, indistinguishability was observed in the points of agency for the conditions of *PasMovWLDelay* and *StAssocPasMov*. In the *PasMovWGrdDec* condition, even with a negative delay of -12 ± 2.4 ms ($M \pm SD$), the average score was 5 ± 1.85 ($M \pm SD$).

The peak-to-peak amplitude of the somatosensory ERP was greater in the condition with a constant low delay (*PasMovWLDelay*), compared to the condition of random delay associated with the stimulus (*StAssocPasMov*). It could turn out that the constant but very small (9 ± 2 ms (M ± SD)) passive movement delay in the *PasMovWLDelay* condition was subjectively similar in its surprise element to the situation of a random variation of the passive movement delay in the *StAssocPasMov* condition. The variation of the passive movement delay in the *StAssocPasMov* condition was organized using the same time distribution as for the target stimulus (see Fig. 2). In addition, the passive movement, despite the randomness of the delay, almost always (in 98% of cases) occurred after the occurrence of the target stimulus.

The behavioral results obtained may be related either to psychophysiological adaptation to new temporary relationships between the targeted stimulus and the intended action, or to insufficient reflection of subjective sensations in the questionnaire's score descriptions, or to the effect of prejudice. The experiment assumed an unambiguous connection between the appearance of the target stimulus and the subsequent action. Such "context pressure" may promote JoA, according to the postdictive model of the influence of prejudice and situational factors. It is tempting to accept as one's own an action that is habitual and necessary, even if it is not felt in this way (no FoA). Probably prejudice had a greater effect on the agency of the motor event than did temporary uncertainty.

It may be impossible to make a correct subjective score-based evaluation of agency under certain conditions, including at least some of those used in our study. Nevertheless, the test conditions in the first part of the experiment, creating situations of high unpredictability (*StIndepPasMov*) and full control over passive movements (*DelegatedMov*) showed plausible scores, which indicate, at least, the participants' understanding of the instruction and the adequacy of the scale for its extreme values. Finally, the outcome of relatively high agency scores for the passive movement with negative "reaction time" can be explained as follows. Due to the servo drive's adjustment to individual motion parameters, its movement was not very fast and partially overlapped with the time of the target stimulus. Thus, behavioral assessment in this situation can make a subjective determination of the movement's beginning impossible, although it could in general be recognized as self-initiated, again, in view of the need to mentally respond to the target stimulus.

Conclusions

The results obtained in this study support the idea of the independence of the sensory (FoA) and the evaluative (JoA) aspects of the agency experience (Gallagher, 2000, 2006; Synofzik et al., 2008). In part, they can be explained using a predictive feed-forward model (Blakemore et al., 1999, 2000, 2002), since in the absence of active motions, the FoA effect was not observed. On the other hand, the negative relationship between the amplitude of somatosensory ERP and the delay of the passive motion, as well as the dependence of the ERP amplitude on the surprise factor of passive motion, allows the following, more elaborated explanation. Prediction of the movement's sensory effects remains possible with passive movements, provided that it is coordinated with the mental representation of the situation. If so, then the feed-forward model is not applicable in this case, since the model requires active implementation of motor commands to construct predictions of sensory consequences. This argument, together with the presence of JoA in the context of various delays between the target stimulus and the motor event, even when these delays are extremely short, provides crucial evidence in favor of postdictive models of agency experience, in which the leading role is assigned to the situational factors of prejudice and contextual knowledge related to the action (Wegner & Wheatley, 1999; Wegner et al., 2004). It seems that the "context pressure" of the experimental situation, which presupposes a mandatory response to the stimulus, enables prediction of passive movements and their sensory consequences.

In future studies, the subjective evaluation of the time intervals of the action-result (e.g., Engbert, Wohlschläger, & Haggard, 2008) can be used as an alternative and possibly a more sensitive measure than the survey-based agency estimation. This method follows from the "intentional binding" effect discovered by Haggard, Clark, & Kalogeras, (2002): the interval between the action and the subsequent event is estimated shorter if it is accompanied by a sense of agency.

Acknowledgments

This work was in part supported by the Russian Science Foundation, grant RScF, project No. 14-28-00234.

References

Abbruzzese, G., Ratto, S., Favale, E., & Abbruzzese, M. (1981). Proprioceptive modulation of somatosensory evoked potentials during active or passive finger movements in man. *Journal of Neurology, Neurosurgery & Psychiatry*, *44*(10), 942–949. doi: 10.1136/jnnp.44.10.942

Bayne, T., & Pacherie, E. (2007). Narrators and comparators: The architecture of agentive self-awareness. *Synthese*, *159*(3), 475–491. doi: 10.1007/s11229-007-9239-9

Bays, P.M., Wolpert, D.M., & Flanagan, J.R. (2005). Perception of the consequences of self-action is temporally tuned and event driven. *Current Biology*, *15*(12), 1125–1128. doi: 10.1016/j.cub.2005.05.023

Bernier, P.M., Burle, B., Vidal, F., Hasbroucq, T., & Blouin, J. (2009). Direct evidence for cortical suppression of somatosensory afferents during visuomotor adaptation. *Cerebral Cortex*, *19*(9), 2106–2113. 10.1093/cercor/bhn233

Benazet, M., Thénault, F., Whittingstall, K., & Bernier, P. M. (2016). Attenuation of visual reafferent signals in the parietal cortex during voluntary movement. *Journal of Neurophysiology*, *116*(4), 1831–1839. doi: 10.1152/jn.00231.2016

Blakemore, S.J., Wolpert, D.M., & Frith, C.D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience, 1*(7), 635–640. doi: 10.1038/2870

Blakemore, S.J., Frith, C.D., & Wolpert, D.M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience*, *11*(5), 551–559. doi: 10.1162/089892999563607

Blakemore, S.J., Wolpert, D.M., & Frith, C.D. (2000). Why can't you tickle yourself? *Neuroreport*, *11*(11), R11–R16. doi: 10.1097/00001756-200008030-00002

Blakemore, S-J, Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, *6*(6), 237–242. doi: 10.1016/S1364-6613(02)01907-1

Desantis A., Roussel C., & Waszak F. (2011). On the influence of causal beliefs on the feeling of agency. *Consciousness and Cognition, 20*, 1211–1220. doi: 10.1016/j.concog.2011.02.012

Eimer, M., Maravita, A., Van Velzen, J., Husain, M., & Driver, J. (2002). The electrophysiology of tactile extinction: ERP correlates of unconscious somatosensory processing. *Neuropsychologia*, *40*(13), 2438–2447. doi: 10.1016/S0028-3932(02)00079-9

Engbert, K., Wohlschläger, A., & Haggard, P. (2008). Who is causing what? The sense of agency is relational and efferent-triggered. *Cognition, 107*(2), 693–704. doi: 10.1016/j.cognition.2007.07.021

Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, *4*(1), 14-21. doi: 10.1016/S1364-6613(99)01417-5

Gallagher, S. (2002). Experimenting with introspection. *Trends in Cognitive Sciences, 6*, 374–375. doi :10.1016/s1364-6613(02)01979-4

Gallagher, S. (2006). *How the body shapes the mind*. Oxford: Oxford University Press.

Gentsch, A., & Schütz-Bosbach, S. (2011). I did it: Unconscious expectation of sensory consequences modulates the experience of self-agency and its functional signature. *Journal of Cognitive Neuroscience*, *23*(12), 3817–3828. doi: 10.1162/jocn_a_00012

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382–385. doi: 10.1038/nn827

Hesse, M.D., Nishitani, N., Fink, G.R., Jousmäki, V., & Hari, R. (2010). Attenuation of somatosensory responses to self-produced tactile stimulation. *Cerebral Cortex*, *20*(2), 425–432. doi: 10.1093/cercor/bhp110

Karniel, A. (2002). Three creatures named 'forward model'. *Neural Networks*, *15*(3), 305–307. doi: 10.1016/S0893-6080(02)00020-5

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22*(140), 5–55.

Kumar D., & Srinivasan N. (2014). Naturalizing sense of agency with a hierarchical event-control approach. *PLoS One, 9*, e92431. doi: 10.1371/journal.pone.0092431

Kumar, D., & Srinivasan, N. (2017). Multi-scale control influences sense of agency: Investigating intentional binding using event-control approach. *Consciousness and Cognition*, *49*, 1–14. doi: 10.1016/j.concog.2016.12.014

Moore J.W., Wegner D.M., & Haggard P. (2009). Modulating the sense of agency with external cues. *Consciousness and Cognition*, *18*, 1056–1064. doi: 10.1016/j.concog.2009.05.004

Moore, J. W. (2016). What is the sense of agency and why does it matter? *Frontiers in Psychology*, *7*. doi: 10.3389/fpsyg.2016.01272

Shergill, S.S., White, T.P., Joyce, D.W., Bays, P.M., Wolpert, D.M., & Frith, C.D. (2013). Modulation of somatosensory processing by action. *Neuroimage*, *70*, 356–362. doi: 10.1016/j.neuroimage.2012.12.043

Sidarus, N., Vuorre, M., & Haggard, P. (2017). How action selection influences the sense of agency: An ERP study. *NeuroImage*, *150*, 1–13. doi: 10.1016/j.neuroimage.2017.02.015

Shimojo, S. (2014). Postdiction: Iits implications on visual awareness, hindsight, and sense of agency. *Frontiers in Psychology, 5*: 196. doi: 10.3389/fpsyg.2014.00196

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, *17*(1), 219–239. doi: 10.1016/j.concog.2007.03.010

Tsakiris, M., Schütz-Bosbach, S., & Gallagher, S. (2007). On agency and body-ownership: phenomenological and neurocognitive reflections. *Consciousness and Cognition*, *16*(3), 645–660. 10.1016/j.concog.2007.05.012

Wegner, D.M. & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, *54*(7), 480–492. doi: 10.1037/0003-066X.54.7.480

Wegner, D.M., Sparrow, B., & Winerman, L. (2004). Vicarious agency: Experiencing control over the movements of others. *Journal of Personality and Social Psychology*, *86*(6), 838. doi: 10.1037/0022-3514.86.6.838

Weiskrantz, L., Elliott, J., & Darlington, C. (1971). Preliminary observations of tickling oneself. *Nature*, *23*, 598–599. doi: 10.1038/230598a0

Wolpert, D.M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences*, *2*(9), 338–347. doi: 10.1016/S1364-6613(98)01221-2

Original manuscript received June 12, 2017
Revised manuscript accepted July 01, 2017
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society
Lomonosov Moscow State University

A computational model of consciousness for artificial emotional agents

Artemy A. Kotov

Kurchatov Institute National Research Center, Moscow, Russia
Russian State University for the Humanities, Moscow, Russia

Corresponding author. E-mail: kotov_aa@nrcki.ru

Background. The structure of consciousness has long been a cornerstone problem in the cognitive sciences. Recently it took on applied significance in the design of computer agents and mobile robots. This problem can thus be examined from perspectives of philosophy, neuropsychology, and computer modeling.

Objective. In the present paper, we address the problem of the computational model of consciousness by designing computer agents aimed at simulating "speech understanding" and irony. Further, we look for a "minimal architecture" that is able to mimic the effects of consciousness in computing systems.

Method. For the base architecture, we used a software agent, which was programmed to operate with scripts (productions or inferences), to process incoming texts (or events) by extracting their semantic representations, and to select relevant reactions.

Results. It is shown that the agent can simulate speech irony by replacing a direct aggressive behavior with a positive sarcastic utterance. This is achieved by balancing between several scripts available to the agent. We suggest that the extension of this scheme may serve as a minimal architecture of consciousness, wherein the agent distinguishes own representations and potential cognitive representations of other agents. Within this architecture, there are two stages of processing. First, the agent activates several scripts by placing their if-statements or actions (inferences) within a processing scope. Second, the agent differentiates the scripts depending on their activation by another script. This multilevel scheme allows the agent to simulate imaginary situations, one's own imaginary actions, and imaginary actions of other agents, i.e. the agent demonstrates features considered essential for conscious agents in the philosophy of mind and cognitive psychology.

Conclusion. Our computer systems for understanding speech and simulation of irony can serve as a basis for further modeling of the effects of consciousness.

Keywords: cognitive architectures, psychophysiological problem, theory of consciousness, emotional computer agents, machine humor, simulation of irony, text comprehension

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0304
http://psychologyinrussia.com

Introduction

For many authors, the notion of "consciousness" is not a strong scientific concept, but rather an element of a "naïve world picture" (e.g., Bulygina & Shmelev, 1997). Usually consciousness is described as a subjective space, which holds mental processes (percepts, representations, thoughts), and is available for observation in the same way as the surrounding physical world. Often a person refers to this space as their "self" ("me" or "I"), although this notion can also refer to an internal world in a wider sense: to one's own knowledge, principles, and, of course — one's own body. "Self" (or consciousness) is also considered to be a source of voluntary actions. When a person acts automatically or reflexively, it is usually suggested that not only the stimulus but the reaction itself resides beyond the boundaries of consciousness, as if something external imposes the reaction on people. However, if a person acts "rationally" or "deliberatively", it is considered that consciousness is the source of these acts.

The "naïve view" of a person with regard to his or her own emotions is more complicated. On the one hand, people usually attribute their own emotional actions to rational choice. Metaphorically, emotions are seen to constitute an external force which leads a person to execute a certain action. On the other hand, one usually highly evaluates one's own emotions, according them priority with regard to any ensuing choices and the evaluation thereof. One can even say: *I understand, but intuitively **I** feel different* or *I know what I ought to do, but **I** want to do something different*. In these cases one addresses one's own emotions as a the "true Self". At the same time, the machinery of emotions is not consistent. M. Minsky (1988, p. 165) introduced the concept of "proto-specialist" — a simple model of emotions and drives for an "artificial animal". Each proto-specialist is responsible for the detection of a dangerous (or lucrative) situation and competes with other proto-specialists in order to force the body (the whole organism) to execute a suggested action. The balance between proto-specialists (or other mental agents) will constitute the central point for our further study of consciousness.

As the notion of consciousness is subjectively evident but hard to address scientifically, numerous approaches to this problem have emerged (see, e.g., Chernigovskaya, 2016; Velichkovsky, 2015). In the philosophy of mind, the notion is linked to studies of *understanding* — consciousness is frequently considered as an "organ" for understanding: a mental processor or container for the understood meaning. J. Searle (1980), in his "Chinese room argument", examined and criticized a theoretical design of an *understanding* computer agent — a digital computer. Searle argued that a digital computer solely operates with the data, following the defined rules, and implied that the whole model (and any computer) could not achieve the skill of understanding. In a similar way, T. Nagel (1974) argued that consciousness is incognizable, as no technology (imitation or physical transformation) can let us know what it is like to be a living being — *a bat*. Modern approaches shift the emphasis in this classic discussion: computer models of understanding (speech processors, robot behavior planners) are designed to operate with texts or with behavioral patterns. They are not intended to "make us feel like a robot" nor to "show us the modeled consciousness". So the model of consciousness cannot be falsified, if it does not immerse us in the modeled consciousness, just as engineering models are not designed to "make us feel like a bridge", but rather to test the bridge in different situations.

Another major approach to consciousness is the attempt to describe introspection or self-awareness. It is suggested that introspection is either essential for consciousness or is a form of consciousness and thus the simulation of introspection may give us a clue to the simulation of consciousness. An analysis of these theories was recently conducted by M. Overgaard and J. Mogensen (2017). A theoretical model of introspection usually has a "double-layer" architecture, where the first layer is responsible for general cognitive tasks, and the second layer monitors or alters the first layer. In a procedural approach undertaken by A. Valitutti and G. Trautteur (2017), it is suggested that on the first level, a system runs general cognitive tasks, while the second level may inspect and alter these basic operations. An example is a software interpreter, which executes the code, simulates the execution (traces and mirrors the code), and may insert additional instructions based on the examination of a single instruction (*local introspection*), as well as on the entire target program (*global procedural introspection*) (Valitutti & Trautteur, 2017). Half a century ago, M. Minsky (1968) proposed that a living being (a man — M) may have a model of self, M*, which answers questions like "how tall am I?" — and a higher level model, M**, with descriptive statements about M*. Minsky suggested that the distinction between M* and M** leads to a "body and mind" paradox, whereby one cannot explain the interaction between cognition and the brain — as mental and physical structures are natively represented by different models.

Although the "double-layer" architecture is widely used in theoretical studies and computer simulations, the definition of introspection as an essential attribute of consciousness may limit the model: subjectively we may be "conscious" when acting in the real world and thinking about real objects — not only at a time of introspection. Therefore, the model should be elaborated to suggest the state of consciousness in different situations, not only in the state of self-awareness.

In psychology, consciousness is frequently explained by the notion of *short-term memory*. It is suggested that short-term memory is the machinery supporting the mental structures which we subjectively perceive to be the content of consciousness. The computer metaphor, applied here to living creatures, indicates the amount of information (objects, features etc.) that can be simultaneously preserved and processed by the subject (for details of this concept, see B.B. Velichkovsky, 2017). It might be that "simple" creatures have a limited *memory*, reducing their behavior to simple reactions. On the other hand, humans have an extended *memory,* allowing them to operate with language structures, mental images, logical inferences, etc. Adherence to this latter metaphor brings us to some questionable results. Modern computers have a huge amount of RAM accessible by software. This however does not evolve them to a threshold of gradual emergence of consciousness, suggesting that the mode of operation may be far more significant than the amount of data processed.

Theoretical approach

Following the analysis of the "naïve" notion of consciousness, we may define a list of features to be modeled by software to produce a "conscious" agent (if a model of consciousness is indeed possible). A computational model of consciousness should:

- provide space for subjective imaging[1] including pretend images, and establish some kind of coordination between images for further goals;
- generate and verify subjective images or intentions;
- distinguish “self” and “non-self” images, inferences, feelings, or intentions, classify subjective images, and attach subjective feelings to the images;
- handle and possibly solve clashes between conflicting images, feelings, or intentions;
- generate and coordinate actions in a sophisticated way — not only on the basis of pure reactions, but with the consideration of many significant factors

We rely on a cognitive architecture, developed within the Cognition and Affect project (CogAff). This is a “shallow” cognitive model, designed to depict basic cognitive and emotional functions and to be implemented by virtual computer agents (Sloman, 2001; Sloman & Chrisley, 2003). CogAff architecture relies on a “triple-tower” model by Nilsson with a perception module receiving data from the environment, a central processor, and an action unit responsible for the generation of actions (Nilsson, 1988). On different levels, CogAff distinguishes: (a) procedures for emotional processing — *alarms* or *reactions*, (b) deliberative reasoning — models for rational inferences, and (c) models for reflective processes on a “meta-management” level. Entities on each level compete in processing information and in generating output. In CogAff architecture, a lower level of emotional reactions is separated from rational processing by an attention filter. Processes under the attention filter are executed automatically. They can stay removed from attention and consciousness, only to inform the deliberative processing level that a certain reaction took place. At the same time, cognitive structures above the attention filter belong to the deliberative reasoning (or meta-management) levels and simulate the reasoning process of human consciousness. CogAff agents effectively handle some important tasks, like solving conflicts between emotional and rational processes. The architecture also suggests the concept of “tertiary emotions”, which use meta-management to inject mental images that have originally driven an emotional response — as in the case of phobias and longing — so that the agent frequently returns to the emotional stimulus in its “thoughts”.

With all the advantages of the model, developers can rely on the labels attached to model levels to define “deliberative” processes or “consciousness”. Unfortunately, simple labeling of different levels does not explain the structure of consciousness: if a process operates on the level labeled as “consciousness”, this does not imply that the process is innately *conscious*. Instead, we have to suggest a specific architecture, operating with different mental objects and sufficiently elaborate to represent an “architecture of consciousness”. On the way to the definition of this architecture,

1 “Images” are understood in the present context as visual, auditory, spatial, linguistic, and emotional representations.

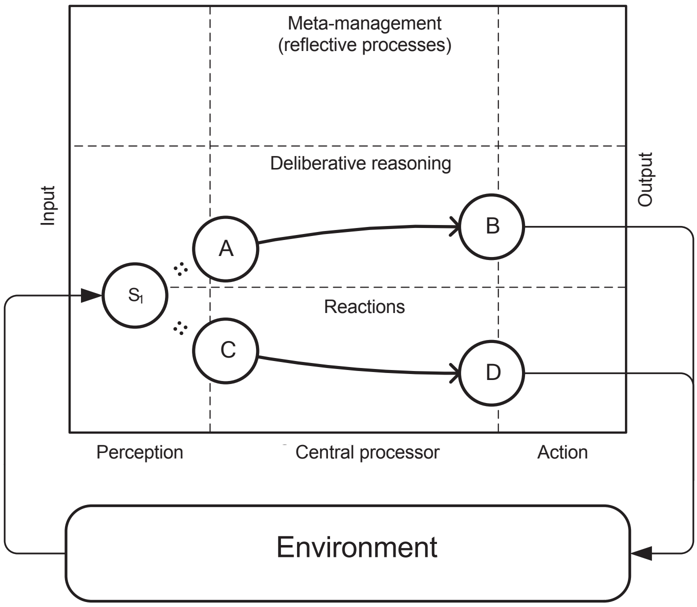


Figure 1. CogAff (Cognition and Affect) scheme as a shallow model of a software cognitive agent

we may suggest several alternatives on how the natural consciousness might be designed. There are the following possible options:

(a) Human consciousness is located in some "spiritual" world and is not connected to any physical (biological) substrate of the body. In this case, all scientific studies of the brain are useless because consciousness cannot be implemented in any hardware or software architecture.

(b) Human consciousness resides in some elements of the brain — molecules, proteins, or other units — and is explained by their physical features. In this case, consciousness cannot be implemented on any hardware, but only on the natural brain tissues or neural network.

(c) Human consciousness is a structural scheme, a mechanism for the interaction of ideal or physical entities. In this case there might be a possibility to implement consciousness with the help of a computer model, relying on existing or future algorithms.

In our view, option (a) does not meet the law of parsimony — even if consciousness has an ideal nature, this option can be preferred only if all conceivable approaches within (b) and (c) are exhausted. Option (b) has an immediate relation to the psychophysiological problem, and suggests that consciousness stems from specific physical (chemical or biological) elements within the brain. If these elements form some structural schemes, suggesting a machinery of consciousness, then these schemes can be modeled by theoretical or real computer architectures — and we arrive at option (c). However, if consciousness is connected to some immanent features of the physical brain (as the feature "golden" is connected to the nature of the mineral "gold"), then we arrive at the paradoxical inference that consciousness is a characteristic of matter. Following these inferences, we choose option (c) as the most substantiated. This option suggests that consciousness is a structural scheme, implemented in the physical machinery of the brain. It can be generally described via a theoretical model and run on data processors with various hardware. This approach also suggests that consciousness (or *the effects of consciousness*) can be studied and modeled even before the "psychophysiological problem" is solved.

If we follow option (c), we should roughly assess the number of elements within this architecture. It is usually expected that computer models of consciousness should simulate physical brain structure, and thus should operate with with the scale of the whole brain and not by a structure with fewer elements. We shall follow the opposite approach, however, and suggest that there does exist a *minimal architecture of consciousness*, which is simpler than that of the entire human brain. In the present publication we shall represent our view of the key features of this minimal architecture. We rely on a theoretical model operating with scripts and its computer implementation suggesting that consciousness or *the effects of consciousness* appear if an agent has the capacity to process one stimulus simultaneously with a number of scripts, and if a subsequent script during its activation can access a set of scripts at the previous level. A key example of our approach is the computer simulation of irony.

The model of consciousness

In many approaches it is suggested that an emotional analysis of input competes with rational (conscious) processing. So the procedure of emotional text processing may be a key to the understanding of the architecture of consciousness. Earlier, we (Kotov, 2003) presented a list of *dominant scripts* (*d-scripts*), responsible for the recognition of emotional patterns in a natural text, and competing with rational procedures (*r-scripts*) during input processing. A script is a sort of production (inference) with an if-statement — *initial model* and action — *final model.* The list of negative d-scripts consists of 13 units responsible for the recognition of patterns: *It affects your health; They will kill you* (DANGER d-script); *There is no way to go* (LIMIT); *They are just crazy* (INADEQ); *Nobody needs you* (UNNEED); *Everything is useless* (FRUSTR), etc. These scripts appear in dialogues involving conflict (*You don't even care if I die!*) and in negative propaganda (*The government does not care!*). The list of positive scripts includes 21 units for: *It is beautiful* (VIEW); *This sofa is so nice and cozy* (COMFORT); *You control the situation perfectly* (CONTROL); *Everybody loves you* (ATTENTION), etc. These scripts appear in compliments, advertising, and positive propaganda.

A computer agent operated by *d/r-scripts* proved to be capable of simulating speech irony (Kotov, 2009). The agent acted in the following way: when receiving input about an event such as "Someone is hitting you", it activated a negative script DANGER and was ready to reply, *I was hit! You — idiot!* However the agent was suppressing the direct expression of DANGER script in speech; instead, it was looking for a positive script with the highest level of activation — this was the ATTENTION script, usually expressed in the utterances *It's a good thing you have paid*

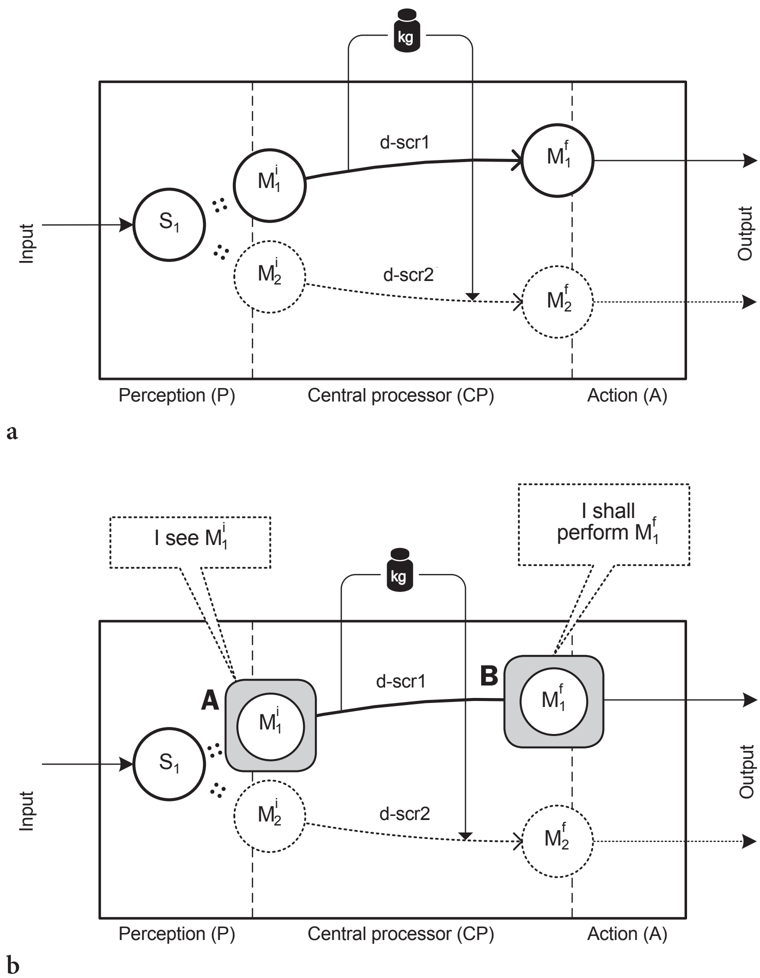


Figure 2. Architecture of an agent inhibiting alternative reactions

attention to me! It's a good thing you care about me! The agent used the utterances from ATTENTION to express the concealed activation of the DANGER script, adding to the utterances a marker of irony (see further details in Figure 5).

A balance between different scripts forms the cornerstone of our approach to the *minimal architecture of consciousness*. Let us see how this balance is achieved during processing of a stimulus in simple reactive architectures, having no relation to conscious processing (Figure 2). A stimulus S_1 may activate a number of scripts — in particular d-scr1 with high activation, and d-scr2 with lower activation. In Figure 2a we demonstrate an architecture that selects the winning script through script displacement (inhibition of scripts with lower activation). If d-scr1 has received higher activation, then an alternative d-scr2 is suppressed and never appears in the output (indicated by a dotted line).

Quite frequently the notion of consciousness is explained through the notions of *operative memory* and *attention*. We shall use the term *scope of processing* (or *processing scope*) in a similar sense. We affirm that the processing scope may contain initial and final models of the scripts. We can compare the processing scope to a desktop with work materials: in order to add any new material, we have to clear space on the desk and remove some older papers. Any inference can be made only on the basis of materials already on the desktop. All papers once removed from the desk no longer exist and are not accessible for immediate cognitive operations. We note that the processing scope of a simple agent (as in Figure 2) contains only one script model. Then, for the agent in Figure 2, the processing scope initially will be

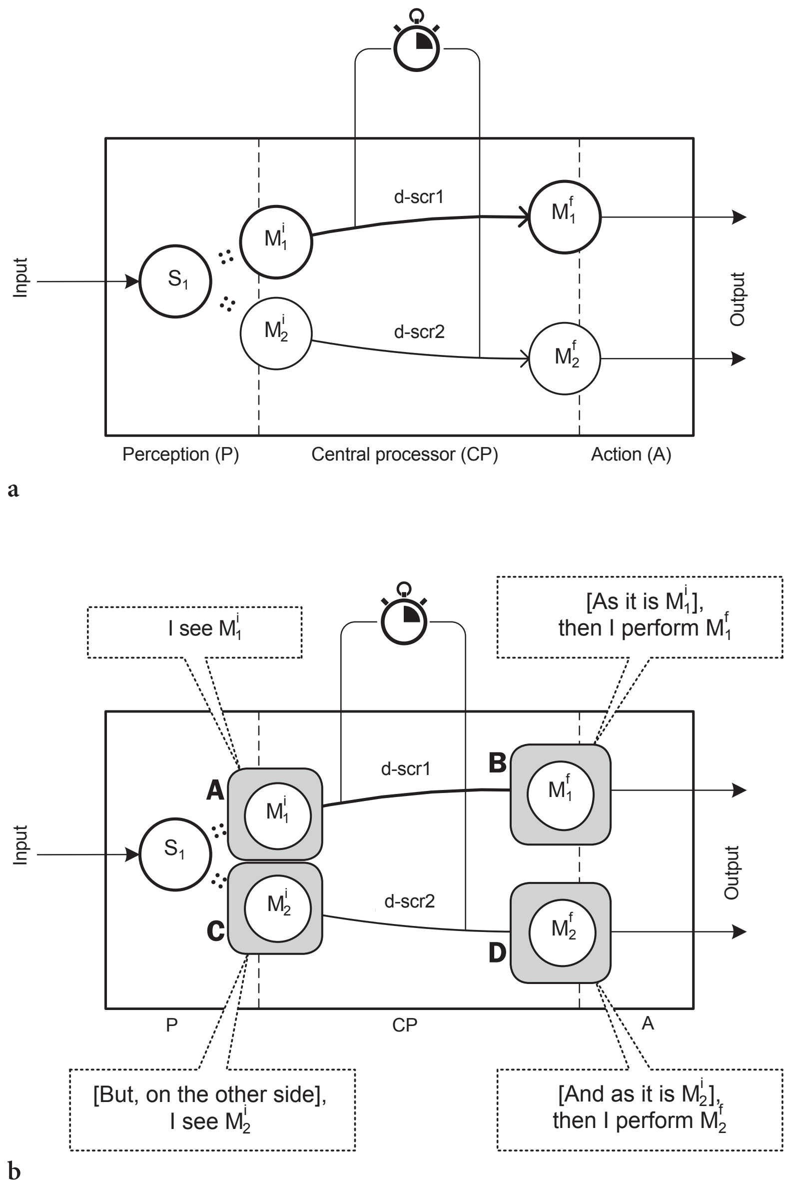


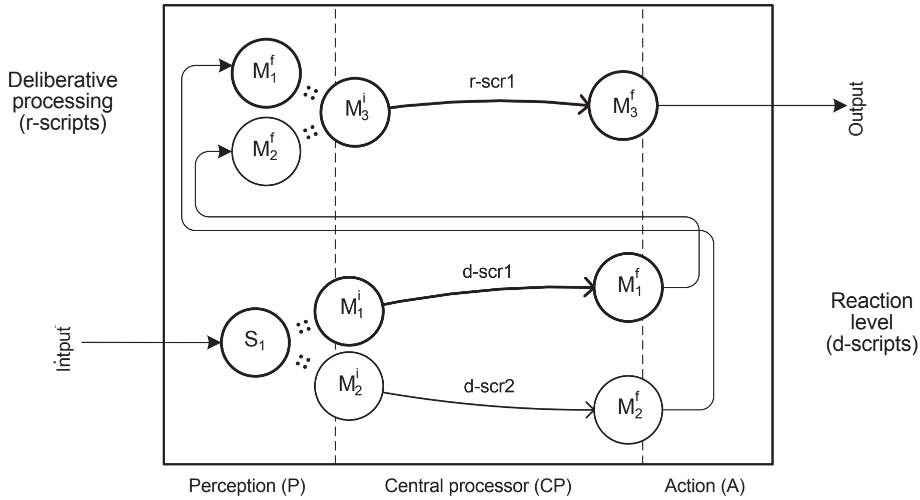
Figure 3. Architecture of the agent with temporal distribution of scripts

in position A, and contain M_1^i model (Figure 2b) — at this stage the agent interprets a stimulus S_1 as M_1^i and believes that M_1^i takes place in the reality in front of him. While activating script d-scr1 and moving to the final model of the script, the agent replaces the contents of the processing scope: in B position the whole processing scope is occupied by the M_1^f model. If, for example, the initial model M_1^i had the content "Somebody is hitting me", then the final model M_1^f may provoke the responsive aggression of the agent. An alternative script d-scr2 will then be inhibited and will never be used to react to the S_1 situation.

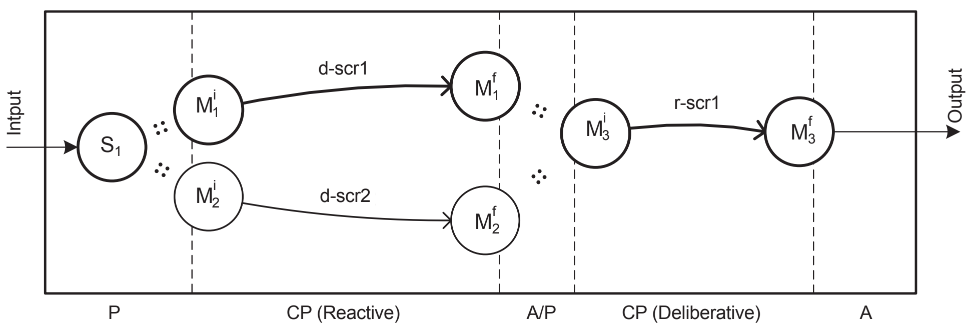
The reactions of an agent may be distributed in time (Figure 3). In this situation, the input S_1 will activate scripts d-scr1 and d-scr2 (as in the previous case). First, d-scr1 will take place, as a script with higher activation, while d-scr2 will be temporarily suppressed. Second, d-scr2 will take place after a standby period. For example, if we "step on the foot" of the agent, he may, first, curse, and second, suggest a socially acceptable reply, saying, *It's all right!* We used the temporal distribution of scripts to simulate the spoken emotional behavior of a computer agent (A. Kotov, 2007). In this architecture, the processing scope will sequentially reside in the A, B, C, and D positions — Figure 3b. First the agent interprets S_1 as $A(M_1^i)$. The interpretation M_2^i at this moment is also constructed by the agent, but this is temporally delayed and no longer remains within the processing scope. Then, the agent reacts to M_1^i — moving to B position and executing actions as defined by M_1^f. When the d-scr1 script is completely processed, the agent shifts to d-scr2. Now it moves representation M_2^i to the processing scope (position C) and then proceeds along with d-scr2 to the inferences or actions of M_2^f (position D). The agent may lack the resources to discover the co-reference of $A(M_1^i)$ и $C(M_2^i)$ so as to understand that these are two different representations of the same situation S_1. If the processing scope presents only one model, then M_1^i и M_2^i will never appear at the processing scope simultaneously so as to be compared by the agent — and the agent will not discover their partial similarity and co-references. As for the result of this limitation, the agent may construct contradictory representations of one and the same situation — and react accordingly to these representations.

Agents shown in in Figures 2 and 3 have very simple architectures: they use scripts from only one level of processing (d-scripts) and can place at the processing scope only one model. More sophisticated agents combine the reactive level with deliberative processing and can activate both d-scripts and r-scripts. In CogAff architecture, this situation can be represented as seen in Figure 4a: final models M_1^f and M_2^f of d-scripts d-scr1 and d-scr2 from the action component of the reaction level are transferred to the input of the deliberative processing level, and may activate a rational script — r-scr1.

We shall rearrange this scheme and draw the processing cycle as a straight line (Figure 4b). Let the scripts d-scr1 and d-scr2 reside on the left from r-scr1, while at the same time keeping in mind that they belong to two different levels of processing: reactive and deliberative. R-script r-scr1 can be activated by M_1^i (then M_1^f is interpreted as M_3^i) or by M_2^f (then M_2^f is interpreted as M_3^i). If the r-script is activated by one of these models, then we can get similar architectures with inhibition or temporal distribution of scripts — as we have seen before (Figures 2 and 3). The main difference is that these architectures work on the upper — deliberative — level. However, we have to pay attention not to the sequential processing, but to



a



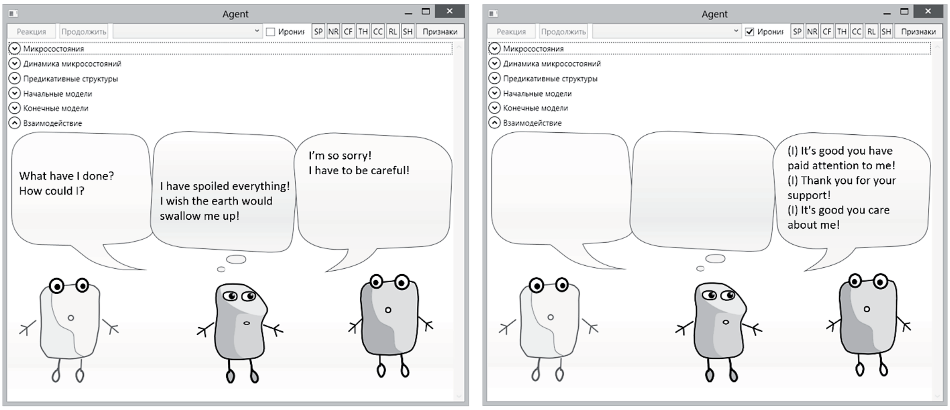
b

Figure 4. Two-level architectures, distinguishing reactive and deliberative processing levels

the simultaneous processing of competing scripts. Consider architectures where scripts and procedures on upper levels have access to several scripts activated on a previous level. In particular, such a mechanism provides a machinery for irony and ironic replies. Irony for us constitutes a significant example, as it is usually considered to be a sophisticated cognitive task, requiring strong conscious processing.

Computer model of irony

Earlier, we represented a computer agent simulating irony with the help of d/r-scripts (Kotov, 2009). In Figure 5a we show an interface where a Green computer agent (at the center) interacts with other agents: Yellow (on the left) and Grey (on the right). Green receives different predicative structures at its input — these can be system events generated by certain system states, interaction with a user (e.g., by mouse clicks) or semantic components constructed by a syntactic parser as a result of natural text analysis. In a case of ironic behavior (Figure 5b), the agent receives an event “Green (other) is hitting Green (self)”, evaluates this event as negative, but suppresses output of curses, and replies ironically: *Thank you for your support!* and *It's a good thing you care about me!* The ironical nature of the text is indicated by the (I) marker in the interface.



a b

Figure 5. Emotional computer agents, software interface

The software processor of the computer agent contains a number of scripts — positive and negative d-scripts responsible for emotional reactions, and r-scripts responsible for rational and socially acceptable replies. The agent compares each incoming event (semantic predication) with the initial models of scripts, calculates the degree of similarity, and defines the activation level for each script. Then the scripts are sorted by the degree of activation. The most activated scripts obtain control over the agent: the agent will then perform gestures and output utterances, as defined for that script.

Following Table 1, d-script DANGER gets the highest activation, 4.1097, after processing the "Somebody hit me" event. If this script gets control over the (Green) agent, the agent complains and swears that "He has been beaten", or shouts at the counterpart Grey agent. Instilling irony, the agent suppresses direct expression of the winning negative script and chooses a positive script with the highest activation. As seen in Table 1, these scripts are CARE (5th line), ATTENTION (6, 12, and 13th lines) and COMFORT (15th line). They all are accorded a similar degree of activation, 2.3482, almost twice as low as that of DANGER (4.1097). From the point of view of the agent, — DANGER is the most relevant classifier (script) for the situation "Somebody hit me"; however, the agent has the ability to choose a positive script with the highest activation to output an ironic answer. ATTENTION type 1 with output utterances *Thank you for your support!* and *It's a good thing you care about me!* was among others selected by the agent in our first experiments. Following the activation level, ATTENTION is not a relevant class (script) for the initial stimulus and can be used only as an extension or a substitute to express DANGER. The initial model of the ATTENTION script is not "what actually takes place" (because some "danger" takes place) and not "what the agent actually feels" (because the agent feels the "danger" — "fear" or "aggression"). Yet this classification of the initial stimulus is still preserved and may be used in a

Table 1. Activation of scripts for an event "Somebody is hitting me"

No.	Score	Script	Possible output
1	4.1097	DANGER	*You will kill me!*
2	3.3482	LIMIT	*You limit me!*
3	3.3482	SUBJECT	*You like to command!*
4	2.3482	PLAN	*You meant that!*
5	2.3482	CARE	*You care about me!*
6	2.3482	ATTENTION type 1	*It's good you have paid attention to me!* *Thank you for your support!* *It's a good thing you care about me!*
7	2.3482	RULES type 1	*It is all right!*
8	2.3482	RULES type 2	*What shall I do in return?*
9	2.3482	Reconciliation	*It is for the best!*
10	2.3482	INADEQ type 4	*You are an idiot!*
11	2.3482	INADEQ type 5	*You don't know what you are doing!*
12	2.3482	ATTENTION type 2	*You are great!*
13	2.3482	ATTENTION type 3	*You understand me!*
14	2.3482	DECEIT	*You lie to me!*
15	2.3482	COMFORT	*I feel great!*
16	2.0133	EMOT	*You are hysterical!*
17	2.0133	SUBJECT type 1	*You think only about yourself!*

communication. This is possibly because the ATTENTION script was not inhibited by other scripts, and the mechanism of irony could access this script among other possible reactions. It means that the processing scope should have some minimal size (here — 15 scripts), to contain a list of scripts with similar or higher activation compared to those suited for the activation of ATTENTION type 1; this ensures that the mechanism of irony can select a suitable "ironic" reaction from the processing scope. If the processing scope contains a number of scripts with different levels of activation, then the agent may differentiate the scripts as "more/less relevant to the situation" or as "my own reactions"/"possible reactions" — this choice can be made simply by means of the activation level. In previous architectures (Figures 2, 3), script activation itself indicated the relevance of the script. Alternatives with lower activation (less relevant scripts) were inhibited or delayed. The agent did not have to compare scripts depending on their activation — this function was effectively executed by an inhibiting process or by a timer. However in the case of irony represented here, the processing scope maintains the ATTENTION type 1 script, which has quite a low activation (not a relevant factor), is neither inhibited nor delayed, and can be accessed by a special communication strategy — making use of irony.

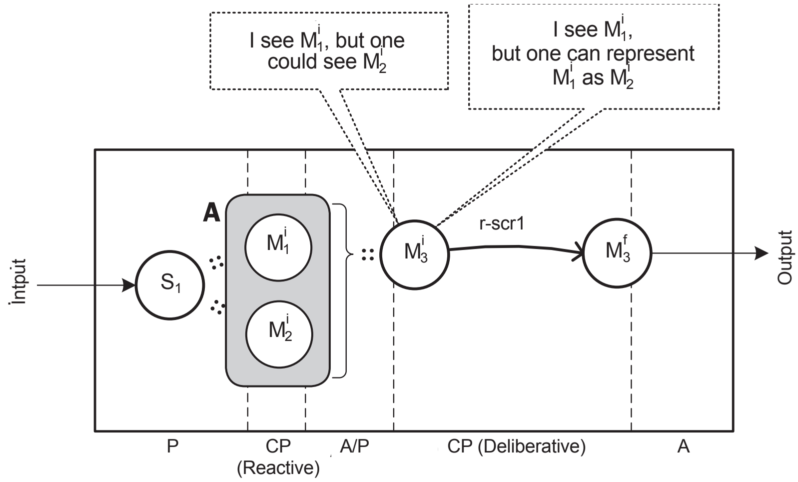
An important feature of the irony mechanism is that it distinguishes (a) a "true" script, which corresponds to the situation and the agent's feelings (in particular, DANGER in the situation of aggression), and (b) an "ironic" script, targeted at the addressee and not reflecting the agent's "true" feelings. Thus the procedure of irony obtains access to two scripts of different degrees reflecting the inner world (or "self") of the speaker and opposite in their evaluation of the situation. Thanks to this architecture, higher-level scripts (or other processing mechanisms, such as irony) can observe the conflict between scripts activated on a lower level and select the scripts that best correspond to the *self* of the speaker. In our example DANGER will better correspond to the speaker's *self* — if we understand *self* as a subjective emotional evaluation, and the ATTENTION script will be targeted at the communication in order to conceal real emotions or to obfuscate the social (communicative) image of the speaker.

In one of the versions of our computer agent we have limited the list of processed scripts to 4 in order to reduce memory load. This change switched off the ability of the agent to synthesize ironic utterances. A negative event could activate several negative scripts, which occupied all 4 slots in the processing scope. In this case the "best" positive script was left out of the allocated memory and could not be accessed through the mechanism of irony. The extension of the processing scope allows the agent to choose the most relevant positive script in a negative situation (and vice versa), even when "top memory slots" are occupied by negative d-scripts, more relevant in a negative situation.

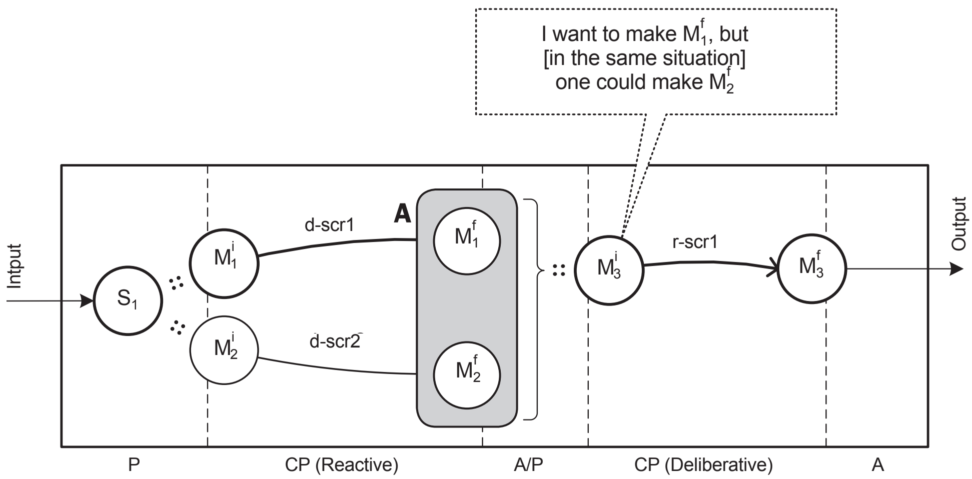
Implication for the model of consciousness

In general, the architecture of irony, and possibly consciousness, requires that: (a) a set of scripts is maintained simultaneously in the processing scope and (b) a further r-script (or mechanism of irony) "sees" these scripts — thus gaining access to many scripts in the processing scope — and is able to distinguish these scripts depending on their activation. This architecture is represented in Figure 6.

Let an incoming stimulus S_1 activate scripts d-scr1 and d-scr2, where both initial models of these scripts are kept in processing scope A. Let us consider the situation whereby an r-script gains access directly to the initial models of these scripts — Figure 6a (final models of this scripts are not shown on the figure). In the processing scope, model M_1^i obtained higher activation, and model M_2^i lower activation. If script r-scr1 can detect this distinction, then the agent "knows" that a situation M_1^i is taking place (the agent "sees" M_1^i); however, in this situation one could see M_2^i. For the agent, it means that it has both "real", as he believes, representation M_1^i and an alternative representation M_2^i (or even a set of such representations). In particular, M_2^i may be used for irony, for the representation "in another situation I could see here M_2^i", "this situation can be represented as M_2^i", "somebody else can see here M_2^i". Thus the extension of processing scope and the ability of r-scr1 to distinguish models in this scope allow the agent to construct a range of "more real" and "more fantastic" representations of an initial S_1 stimulus. The agent thus becomes capable of distinction between reality and alternative representations of reality.



a



b

Figure 6. Architecture of the computer agents for simulation of irony and *the effects of consciousness*

Now consider another situation, where an r-script gets access to the final models of d-scripts — Figure 6b. Here models M_1^f and M_2^f are inferences from an initial situation, actions to be executed by the agent, or goals to be achieved. In any case, M_1^f and M_2^f reflect possible reactions of the agent to the initial stimulus S_1. As in the previous case, the agent may start to distinguish M_1^f and M_2^f as alternative reactions to the situation S_1 if the following conditions are satisfied: (a) processing scope A is big enough to contain M_1^f and M_2^f; (b) r-scr1 has access to both M_1^f models; and the agent can identify the models as alternative reactions to S_1 and at the same time can distinguish the models, based on some differential semantic features. M_1^f obtains higher activation than M_2^f, as d-scr1 initially was more highly activated than d-scr2. While observing the difference in activation level, the agent may conclude that M_1^f is the main reaction to S_1, and M_2^f is an alternative reaction, suitable in the following situations:

- "In a bit different situation *I* could decide/make M_2^f";
- "In a bit different mood/state *I* could decide/make M_2^f";
- "Somebody else in this situation could decide/make M_2^f".

Thus, observing M_1^f and M_2^f, the agent may conclude that some of the available reactions correspond to its *self* (M_1^f), while other reactions are alternatives that less precisely correspond to its *self* (M_2^f) — they can apply to different situations or to different subjects. In other words, in the range M_1^f, M_2^f, …, M_n^f the agent observes the difference between *self* and *non-self* — actions and inferences that the agent attributes to itself, and actions and inferences that the agent has constructed, but does not attribute to itself — that can be only done in other situations or to other people (subjects).

All the represented architectures implement the distribution of alternative scripts. These scripts are not mixed and always choose a "leader", which further controls the agent's performance at each moment. The most important difference of architecture in Figure 6 is that the choice between scripts and their evaluation, is executed by a script at the next processing level, while in the architectures depicted in Figures 2 and 3, the selection of scripts is managed by a mechanism external to the scripts space — an inhibitory process (Figure 2) or temporal distribution (Figure 3). Thus, when moving to the architecture in Figure 6, we observe an "interiorization" of the mechanism for script evaluation and selection. This cognitive evaluation, however, can select not only the most activated script of the previous level — it can take into account other, less activated and less relevant scripts. For example, it can suggest the utterance *It's a good thing you care about me!* as an ironic answer. Less activated scripts can also serve as a matter for imagination ("what could take place", "what I could do") and the theory of mind ("what another person could decide/do").

Scope and limitations of the study

Based on the example of irony, we intended to show that the processes able to explain the architecture of consciousness (demonstrate *the effects of consciousness*) operate at the boundary between the reactive and deliberative processing levels, where an r-script interacts with several activated d-scripts. As we expected, the level of processing does not play the key role here. The same effects can appear during the interaction of d-scripts: "I did M_1^f, but I feel that it is awful and I had to do M_2^f". Similar effects are possible between the deliberative processing and meta-management level, where a person evaluates their own inferences and options for action. So the effects of consciousness are connected with the way scripts interact, not with the location of the scripts in the cognitive model.

We do not claim that our computer model has simulated consciousness or at even the effects of consciousness. We rather consider the software as an illustration of the approach. We have simulated irony as a determined procedure, which always suppresses the most activated negative script and selects a positive script with the highest activation. While moving to the computer simulation of consciousness, it is important to provide more sophisticated interaction between d- and r-scripts.

We do not claim that the random nature of the output or non-determined nature of the processor are important characteristics of consciousness. If a structural scheme of consciousness works on a determined hardware, then for a given input (stimulus S_1) and given the state of the model (scripts), the system will provide one and the same output. At the same time, the represented model has a source of pseudo-random choice: it is evident from Table 1, that at least 12 scripts have the same activation level — 2.3482. Five of these scripts can be used for an ironic answer. What is the main factor of this selection? It can be some minor factors such as the order of the scripts in the database and the sequence of their retrieval. This factor can be determined: each time, for a given stimulus $S_{1,}$ the same ironic answer will be selected for each attempt. At the same time, during the development of the model, the influence of this factor can be reduced: input structures may contain bigger sets of features — S_1 stimuli may differ, reducing the determined nature of the selection. During operation, the system may collect preferences for certain particular scripts, depending on previous choices, or, on the other hand, may avoid repetitive answers. This may appear to be a flexible reaction system, in spite of the deterministic nature of the hardware.

Conclusion

Computer systems for natural speech understanding and the simulation of irony, from our point of view, offer an illustration of an approach which can serve as a basis for further simulation of the effects of consciousness. The mechanism of mutual activation of d/r-scripts (or their analogues) and their interaction in the processing scope can be a cornerstone for the computer model — *the minimal architecture of consciousness*. Within this architecture, the agent should activate several scripts in the first stage of processing, place their if-statements or actions (inferences) within a processing scope, and differentiate the scripts according to their activation by a script of the second stage. This provides an opportunity for the agent to simulate imaginary situations, its own imaginary actions, and the pretended actions of other agents.

Acknowledgements

The research is supported by the Russian Science Foundation (RScF grant 17-78-30029) in the part of modeling semantic representation in the human brain. The design of the emotional computer agents is supported by the Russian Foundation for Basic Research (ofi-m grant 16-29-09601).

References

Bulygina, T.V., & Shmelev, A.D. (1997). *Language conceptualization of the world*. Moscow: Languages of the Slavic Cultures.

Chernigovskata, T.V. (2016). *Cheshire smile of the Schrödinger's cat: Language and consciousness.* Moscow: Languages of the Slavic Cultures.

Kotov, A. (2003). *Mechanisms of speech influence in publicistic mass media texts.* (Ph.D thesis). Russian State Humanitarian Institute, Moscow.

Kotov, A. (2007). Simulating dynamic speech behaviour for virtual agents in emotional situations. In A. Paiva, R. Prada & R. Picard (Eds.), *Affective computing and intelligent interaction* (Vol. 4738, pp. 714–715). Springer / Heidelberg, Berlin.

Kotov, A. (2009). Accounting for irony and emotional oscillation in computer architectures. *Proceedings of International Conference on Affective Computing and Intelligent Interaction ACII 2009* (pp. 506–511). Amsterdam: IEEE. doi: 10.1109/ACII.2009.5349583

Minsky, M. L. (1968). Matter, mind and models. In M. L. Minsky (ed.), *Semantic information processing* (pp. 425–431). Cambridge, MA: MIT Press.

Minsky, M.L. (1988). *The society of mind*. New York, London: Touchstone Book.

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review, 83*(4), 435–450. doi: 10.2307/2183914

Nilsson, N.J. (1988). *Artificial intelligence: A new synthesis*. San Francisco: Morgan Kaufmann.

Overgaard, M., & Mogensen, J. (2017). An integrative view on consciousness and introspection. *Review of Philosophy and Psychology, 8*(1), 129–141. doi: 10.1007/s13164-016-0303-6

Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences, 3*(3), 417–424. doi: 10.1017/S0140525X00005756

Sloman, A. (2001). Beyond shallow models of emotion. *Cognitive Processing, 2*(1), 177–198.

Sloman, A., & Chrisley, R. (2003). Virtual machines and consciousness. *Journal of Consciousness Studies, 10*(4–5), 133–172.

Valitutti, A., & Trautteur, G. (2017). Providing self-aware systems with reflexivity. *ArXiv e-prints:* 1707.08901.

Velichkovsky, B.B. (2017). Consciousness and working memory: Current trends and research perspectives. *Consciousness and Cognition, 55*, 35–45.

Velichkovsky, B.M. (2015). Soznanie [Consciousness]. *Bolshaja Rossijskaja Encyklopedia* [The Great Russian Encyclopedia] (Vol. 30, pp. 623–625). Moscow: The Great Russian Encyclopedia.

Original manuscript received August 25, 2017
Revised manuscript accepted September 11, 2017
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society
Lomonosov Moscow State University

COGNITIVE TASKS AND ABILITIES

The concentric model of human working memory: A validation study using complex span and updating tasks

Boris B. Velichkovsky

Faculty of Psychology, Lomonosov Moscow State University, Moscow, Russia

Corresponding author. E-mail: velitchk@mail.ru

Background. Working memory (WM) seems to be central to most forms of high-level cognition. This fact is fueling the growing interest in studying its structure and functional organization. The influential "concentric model" (Oberauer, 2002) suggests that WM contains a processing component and two storage components with different capacity limitations and sensitivity to interference. There is, to date, only limited support for the concentric model in the research literature, and it is limited to a number of specially designed tasks.

Objective. In the present paper, we attempted to validate the concentric model by testing its major predictions using complex span and updating tasks in a number of experimental paradigms.

Method. The model predictions were tested with the help of review of data obtained primarily in our own experiments in several research domains, including Sternberg's additive factors method; factor structure of WM; serial position effects in WM; and WM performance in a sample with episodic long-term memory deficits.

Results. Predictions generated by the concentric model were shown to hold in all these domains. In addition, several new properties of WM were identified. In particular, we recently found that WM indeed contains a processing component which functions independent of storage components. In turn, the latter were found to form a storage hierarchy which balances fast access to selected items, with the storing of large amounts of potentially relevant information. Processing and storage in WM were found to be dependent on shared cognitive resources which are dynamically allocated between WM components according to actual task requirements. The implications of these findings for the theory of WM are discussed.

Conclusion. The concentric model was shown to be valid with respect to standard WM tasks. The concentric model offers promising research perspectives for the study of higher-order cognition, including underlying neurobiological mechanisms.

Keywords: working memory, concentric model, focus of attention (FA), storage, processing, interference, long-term memory (LTM), serial position effects, complex span task, updating task

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0305
http://psychologyinrussia.com

Introduction

Working memory (WM) is a central component in many theories of cognition. It is a system for on-line storage and processing of information serving the completion of an actual task (Baddeley, 2012). There has been an explosion of research interest in WM over the last decades. This is surely driven by its theoretical prominence, but even more by significant practical links between WM and higher-level cognition. WM has been shown to be strongly related to fluid intelligence (Ackerman et al., 2005), and to the effectiveness of complex activities like foreign language learning, understanding instructions, and control of technical systems (Engle, 2002). WM is also related to academic achievement, and its deficit may be a major cause of learning deficits in school-age children (Gathercole & Alloway, 2008). The proper understanding of WM mechanisms may thus have important practical applications.

The research on the structure and function of WM has long been dominated by Baddeley's multi-component model (Baddeley, 1986). However, studies of individual differences in the limits of WM capacity (Daneman & Carpenter, 1980; Turner & Engle, 1989; Engle, 2002; Shipstead, Harrison, & Engle, 2016) shed a different light on this problem. First, these studies showed a close link between WM capacity and controlled attention. Second, they questioned the fundamental distinction between WM, usually associated with short-term memory (STM), and long-term memory (LTM), suggesting that WM is comprised of elements activated in the LTM. For instance, in the embedded processes theory (Cowan, 1999), a three-layer organization of memory is suggested. The basic layer is formed by the unlimited LTM, from which is selected a subset of activated representation (activated LTM = a-LTM), from which in turn emerges a very limited subset of representations in the focus of attention (FA). Information in the FA can be manipulated and is immune to interference and decay, contrary to that in the a-LTM. Components of WM are characterized by different states of activation — a discovery which gave rise to the notion of activation models of WM.

The most elaborated example of such models is the concentric model by K. Oberauer (Oberauer, 2002). This model extends the embedded processes model by differentiating the FA within the FA proper, and the region of direct access (RDA). While the FA holds only the one cognitive representation which is actually being processed, the RDA is a limited-capacity fast-access system responsible for the maintenance of several representations that are task-relevant, and ready for selection into the FA. The concentric model thus suggests a hierarchy of storage systems with functionally very different components. This hierarchy allows us to make very specific predictions about information transfer and usage within WM. However, the empirical evidence for the concentric models is limited to the study of a specialized WM task (the arithmetic updating task, Oberauer, 2002), which limits this model's generalizability.

In this article, we seek to validate the concentric model with the use of standard WM tasks — complex span tasks and updating tasks — applied to several research domains. Before we present the studies, we elaborate on the concentric model in more detail.

The concentric model

The concentric model makes several basic assumptions about the structure of working memory and the function of its components:

1. The FA is thought to contain a single representation which is being currently processed, and to have the highest level of accessibility. Items are loaded into the FA from the RDA. Items in the a-LTM can be prevented from decay by "loading" them into the FA in a cyclical fashion ("rehearsal").
2. The RDA contains 3–4 items which are thought to be especially relevant for the task at hand. Items can be "off-loaded" to the a-LTM if they are no longer relevant. Items in the RDA are immune to interference, which is the only mechanism to cause forgetting (Oberauer et al., 2012).
3. The a-LTM is potentially unlimited in capacity and contains representations activated over a certain threshold. Representations in the a-LTM may be activated either through their FA, or via activation spreading from other representations in the WM. Representations in the a-LTM can be degraded through interference.

The concentric model suggests that the functional organization of WM is aimed at supporting two major functions–information storage and processing. It also provides a hierarchy of storage systems (RDA and a-LTM), the exact meaning of which is to be clarified. A major problem with the validation of the concentric model is that its validity has only been checked against WM updating tasks which were especially designed for this purpose.

However, today there are several standard ways to measure WM. One is to use typical WM updating tasks like the n-back and the mental counters task (Garavan, 1998; Owen et al., 2005). The other, even more important, approach is to use complex span tasks like the operation span task (Turner & Engle, 1989).

Method

Below, we review some original research on the validation of the concentric models using standard WM tasks and different methodologies. Specifically, we will consider four lines of research: 1) experimental studies of WM structure using S. Sternberg's additive factors paradigm; 2) factor analytic studies of WM structure; 3) a study of serial position effects in WM; and 4) a study of WM structure in a sample with LTM deficits. We will conclude with a general discussion of how our findings relate to the organization of WM.

Results

The structure of WM: Sternberg's additive factors approach

Sternberg's additive factors method (Sternberg, 1984) is a statistical approach to identifying independent processing stages in cognition. It suggests experimental manipulation of several factors, which are assumed to selectively influence a number of hypothetical processing stages. If the factors do not interact statistically (the factors are thus additive), it is concluded that the processing stages associated with them are indeed independent. Extending this logic to the problem of identifying separated components in WM, we suggested that statistical analysis of factors selectively influencing the hypothetical components of WM may reveal the structure of WM (Velichkovsky, 2016; Velichkovsky et al., 2015). Specifically, factor additivity would suggest that the corresponding components can be differentiated, and that they function independently of each other. This logic was applied to the analysis of WM components suggested by the concentric model: the FA, the RDA, and the a-LTM.

In one of our studies (Velichkovsky, Nikonova, & Rumyantsev, 2015), we used the most standard tasks for the assessment of WM functions–the complex span tasks (e.g. Conway et al., 2005). These tasks combine storage of an item set in WM with performing an additional processing task. The main outcome measured by the task is the average probability of reporting the correct item in the correct serial position. Processing task complexity was thought to selectively influence the FA. The between-items interference in the storage task was thought to selectively affect processing in the a-LTM (as the RDA is assumed to be immune to interference in the concentric model). Finally, the WM load (the set size) was manipulated to include between two and six elements to affect both the RDA and LTM. This was plausible since the concentric model assumes a strict limit of 3-4 elements for the RDA (Oberauer, 2002), which means that the a-LTM is used for item storage when this limit is exceeded. The concentric models and the additive factors method allowed two specific predictions to be made about factor interactions in this experimental design:

1. Processing complexity should be independent of both interference and WM load, indicating that the FA differs from the RDA and the a-LTM.
2. WM load and interference should interact, indicating that the RDA and the a-LTM can be differentiated with respect to their sensitivity to interference.

In the study, two span tasks were used–the operation span task (Turner & Engle, 1989) and the parity judgment span task (Lepine et al., 2005). In the operation span task, consonant storage was combined with equation verification. Processing complexity was manipulated by the complexity of the equation verification. Two well-established complexity effects were used: the value effect (verification considered easier if the operands are less than 5), and the odd-even effect (verification considered easier if the parity of the true and displayed answer does not match) (Lemaire & Fayol, 1995). Interference was manipulated by using phonological similarity: consonants were varied according to the number of matching phonological features; those with two overlapping phonological features were considered to interfere to a greater extent (Schweppe et al., 2011). WM load was manipulated by presenting sequences of item sets in ascending and descending order (two to six items and six to two items). It was found that the complexity factor did not interact with either the load or interference factors, as predicted.

It was also found that the load factor interacted with the interference factor, again as predicted. This interaction was driven by the fact that the negative interference effect on storage efficiency was present only for loads over three items.

Both a priori hypotheses were thus supported. Exactly the same results were found in the experiment with the parity judgment task. In both experiments it was also found that increasing the WM load to over 3–4 items led to a significant drop in recall performance.

These results allow us to come to two major conclusions. First, processing of information in WM seems to be independent of storage, as there is no systematic interaction between processing complexity and factors affecting storage. It is tempting to interpret these results in structural terms in that the FA may be considered to be shielded from storage components. Second, these results suggest that storage in WM is realized by two distinct systems. One is a limited-capacity system which is insensitive to interference and provides a reliable storage of items (the RDA, in the terms of the concentric model). Another is a system which is sensitive to interference and less reliable. The second system is involved when there are more items to be maintained than is possible for the RDA to hold.

It is tempting to associate this second storage system with the a-LTM, as the above description closely fits the functional characteristics of the a-LTM, as suggested by the concentric model. There thus is a storage hierarchy in WM; the most accessible item is held in the FA, several items are reliably held in the RDA, and there is much less reliable storage in the a-LTM for the rest of the relevant information. In the Discussion section, we will consider why such a storage hierarchy may have evolved for adaptive purposes. For now, it's sufficient to say that the specific predictions drawn from the concentric model were supported by complex span tasks' data.

In another study we replicated the approach used above with updating tasks, specifically the mental counter task (Miyake et al., 2000) and the n-back task (Owen et al., 2005). Updating tasks require the subject to maintain a set of items in memory, and to dynamically change their content; they are considered prototypical tasks for assessing WM functioning. We manipulated processing complexity, WM load, maintenance duration, and interference.

In the mental counter task the subjects had to count colored figures and react if a specific colored figure was presented forthe third time. The updating thus consisted in incrementally activating a mental counter for a color each time the color was presented. Complexity was manipulated by changing the required increment (+1 in the simple condition, +3 in the complex condition). WM load varied between four (within the RDA limits) and six (over the RDA limits) items. For each stimulus it was also registered how many stimuli were presented since the last presentation of this color (that is, measuring how long the corresponding counter was held in WM).

It was found that the complexity factor did not interact with either the WM load or maintenance duration. It was also found that the WM load interacted with maintenance duration for both accuracy and RT. The duration-dependent error and RT increase was larger for WM loads of six items than for WM loads of four items. These results are in full accord with those obtained for complex span tasks, and extend them by showing that the a-LTM may be also sensitive to time-related decay.

In the n-back task the subjects had to react if the current stimulus matched that presented *n* positions before. Processing complexity was manipulated by changing the identification task (identity or parity match). Interference was manipulated by using either numerically highly distinct digits (selected from 1 to 9) or numerically similar digits (selected from 5 to 9). WM load was manipulated by using 1-, 2, and 3-back conditions. The results matched those obtained for complex span tasks and for the mental counters task. The complexity factor did not interact with either interference or WM load, as the concentric model predicted. Interference and WM load did interact as predicted. To be precise, this interaction, driven by the negative effects of interference, was only observed in the 3-back condition.

In general, the results of the updating task also suggest independence of processing in the FA from information storage, and the existence of two storage components. One component is routinely used for the storage of small amounts of information (about 3–4 items) and is immune to interference and (possibly) time-related decay. The second component is recruited for the storage of larger amounts of information, and is susceptible to interference and time-related decay. These storage components correspond to the RDA and the a-LTM.

The structure of WM: A factor analytic approach

Previous studies used an experimental approach to the study of WM structure. Another approach is to investigate the structure of correlations between different WM tasks. Closely related to Baddeley's WM model, this approach has previously been used to identify modality-specific storage systems within WM (Hale et al., 2011; Giofre et al., 2013), or to differentiate the central executive from storage systems (Kane et al., 2007). We sought to apply this approach to the validation of the concentric model. WM tasks of several types (complex span tasks, continuous span tasks, and updating tasks) were employed in order to sample WM functions with different processing and storage requirements. We were specifically interested in checking whether the correlation structure of different WM tasks allows for the identification of latent factors corresponding to WM components, as suggested by the concentric model. The number and content of latent factors was thus the primary research question. It was also of interest whether WM factors are independent of, or correlate with, each other.

In the study the subjects performed complex span tasks (operation span and counting span, Case et al., 1982) and continuous span tasks (parity judgment span and letter reading span, Lepine et al., 2005), as well as updating tasks (n-back and mental counters task). Continuous span tasks are span tasks with an extremely simplified processing subtask (like parity judgment or reading aloud letters from the native alphabet). This task class has the advantage that its performance does not depend on the ability to solve a complex processing subtask like equation verification. Continuous span tasks, like complex span tasks, were shown to reliably measure WM capacity (Barrouillet & Lepine, 2005). Recall accuracy for the maintenance subtask and processing accuracy and RT were registered for the span tasks, and accuracy (hits) and RT were registered for the updating tasks. Several models differing in the number of factors and factor loadings were devised according to theoretical considerations, and submitted to confirmatory factor analysis to assess model fit (fit indexes computed with the *sem* package in the R statistical computing environment, Fox, 2006).

Specifically, Model 1 contained a single factor (corresponding to the unitary models of WM), Model 2 contained two factors, and Model 3 contained three factors (corresponding to the concentric model). Model fit was assessed via χ^2, CFI, RMSEA, and SRMR fit indexes (see *Table 1* for results).

Table 1. Fit indexes for the structural models (see models' descriptions in the text). Fit index values within admissible range (Brown, 2006) are in bold.

Model	χ^2(df), p	CFI	SRMR	RMSEA
Model 1	62.3(20) <0.01	0.55	0.13	0.21
Model 2	47(18) <0.05	0.69	0.13	0.18
Model 3	**19.6(16) >0.05**	**0.96**	0.12	**0.07**
Model 4	**18.1(15) >0.05**	**0.97**	**0.1**	**0.07**
Model 5	**18(14) >0.05**	**0.96**	0.11	**0.08**

In order to determine the number of factors, we compared Models 1, 2, and 3. According to Table 1, Model 1, with a single WM factor, shows the worst fit. Model 2, with two factors, also inadequately describes the data according to all fit indexes. It is notable that the factors in Model 2 were related significantly and positively ($\beta = 0.56$), suggesting that some additional factors may explain the shared variance. Model 3, with three factors, provided a much better fit with insignificant χ^2 and good CFI and RMSEA values. The SRMR value was somewhat out of range, indicating that further improvement was possible for the model.

Thus the model with three factors clearly outperformed the other models. We further studied the question of the independence of factors. To this end we computed fit indexes for eight models, describing all possible combinations of the presence/absence of relationships between the three factors. Only three of the eight models converged on a solution: Model 3 (assuming total independence of the factors); Model 4 (a positive correlation between Factor 1 and Factor 2, $\beta = 0.21$); and Model 5 (a positive correlation between Factor 1 and Factor 2, and a positive correlation between Factor 1 and Factor 3, $\beta = 0.003$). Table 1 reveals that Model 4 provided a better fit than Models 3 and 5, and was the only model with all fit indexes in the acceptable range (Fig. 1). It is notable that the path coefficient between Factors 1 and 3 is very close to zero in Model 5. If it is set to zero, Model 5 is indistinguishable from Model 4. Thus, Factors 1 and 2 seem to be positively related, while Factor 3 is independent of other factors in the best fitting model.

Factor 1 in Model 4 was loaded by recall accuracy in complex span tasks and continuous span tasks. As these tasks have the heaviest WM storage load (and the FA and the RDA are mostly used for processing and storage of intermediary results, especially in complex span tasks), this factor can be associated with the a-LTM. Factor 2 was loaded by continuous span tasks and accuracy of the updating tasks. Continuous span task have less processing demands, thereby making the use of the RDA for processing less probable. Furthermore, updating tasks also presumably use the RDA for the storage of the small number of items being updated (two items

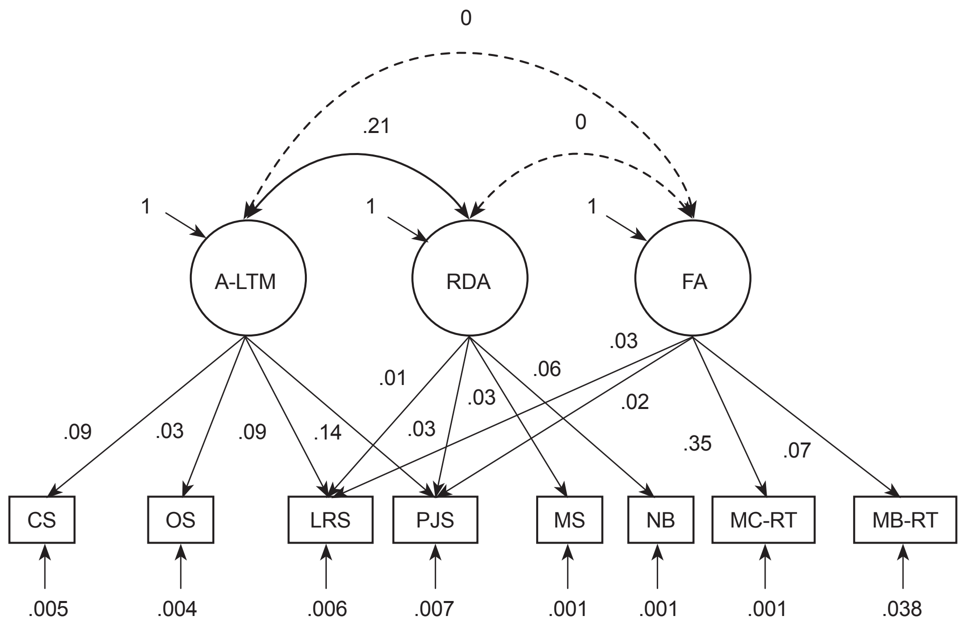


Figure 1. Model 4 (for more details, see the text). CS = counting span, OS = operation span, LRS = letter reading span, PJS = parity judgment span, MC = mental counters (accuracy), NB = n-back (accuracy), MC-RT = mental counters (RT), NB-RT = n-back (RT)

for 2-back and four items for the mental counters version used in this study). Factor 2 may thus be associated with the RDA.

Factor 3 is loaded by the continuous span tasks and the speed of the updating tasks. In interpreting the content of this factor, it should first be noted that continuous span tasks optimize the balance between storage and processing, and thus are indicative of both WM functions. It is also notable that the speed of updating is indicative of more dynamic processing (an updating operation is performed at each stimulus presentation during updating tasks, Botto et al., 2014), while updating accuracy is indicative of more items in storage in WM. For instance, in the n-back task, responses are not guessed only if the nth item is indeed stored in WM. Therefore, Factor 3 can be associated with the FA, the component of WM responsible for information processing.

The three factors in the best-fitting Model 4 correspond to the three WM components suggested by the concentric model: a processing component and two storage components. The storage components differ in the memory load they are able to handle, with the a-LTM handling more information. The a-LTM also seems to be engaged when the RDA is occupied with handling intermediary results for the processing task. These results corroborate the experimental data reported above. For instance, they also show that processing is in a sense independent of storage in the WM, since the FA factor doesn't correlate with the storage factors. On the other hand, both the RDA and a-LTM factors correlate positively, indicating a functional relationship between them. We assume that the relationship between a-LTM and RDA factors may be driven by the presence of information exchange mechanisms between them. It is possible that information is "off-loaded" into the a-LTM from the RDA if it either doesn't fit within the limits of the RDA, or is not considered relevant for the task at hand. Evidence for such "off-loading" was presented by Oberauer (2002) who showed that an item set initially encoded into the RDA can be off-loaded into the a-LTM upon presentation of a cue.

A similar information transfer from the RDA to the a-LTM was also shown with the directed forgetting paradigm (Fawcett & Taylor, 2012). It is perfectly possible that information transfer may occur in the opposite direction (from the a-LTM to the RDA) as previously irrelevant items may be required for the solution of the task at hand. The storage hierarchy in WM may thus be characterized by dynamic information transfer between its components, optimizing the amount of information maintained and its accessibility.

Serial position effects in WM

Serial position effects are observed in immediate recall tasks, and were used in the context of the modal memory model to support the distinction between long-term and short-term memory. These effects can be used to assess the structure of WM and validate the concentric model. In one study, subjects performed the operating span task (a complex span task) and the parity judgment span task (a continuous span task) with WM loads in the 2 to 6 items range. While these tasks are structurally very similar, they differ in the complexity of the processing subtask. In the operating span task, the processing subtask (equation verification) is relatively difficult and requires controlled attention. In the parity judgment span task (parity judgment for a short series of digits) the processing subtask is relatively easy and automatic. We were interested in answering two research questions: 1) whether there are serial position effects during the performance of WM tasks, and 2) whether serial position effects depend on the processing complexity. To this end we assessed both primacy and recency effects based on the relationship between recall accuracy and serial position.

The results suggest that there are pronounced serial position effects during WM tasks (the dependence of recall accuracy on serial position is statistically significant). Items in the middle serial position are poorly recalled. It is tempting to associate the primacy effects with the off-loading of items in excess of the RDA storage limits into the a-LTM (see above), while it is also tempting to associate recency effects with reliable storage of last items in the RDA. Serial position effects are in accord with the storage hierarchy suggested by the concentric model.

However, a more important result is the modulation of serial position effects by the complexity of the processing subtask. First, complex processing leads to a significant reduction of the primacy effect. Second, complex processing also leads to an increased recency effect. In support of these claims, it was found that there was no primacy effect in the operation span task, while there was a significant primacy effect in the parity judgment task. It was also found that the recency effect in the operation span task was significantly higher than in the parity judgment task.

These results suggest that the transfer of information is dependent on the processes and resources also employed by the processing subtask. That is, at least the transfer of information from the RDA to the a-LTM may be dependent on domain-general attention resources. The dependence of storage and processing in WM on general cognitive resources is a common idea in theories of WM (Towse & Hitch, 1995; Vergauwe et al., 2012). Investing resources in an attention-demanding task like equation verification thus prevents information items from being transferred from the RDA to the a-LTM, which negates the primacy effect.

However, the results also suggest that there is a dynamic allocation of cognitive resources within WM with the aim of optimizing both processing and storage efficiency. When the transfer of information into the a-LTM is precluded, storage in the RDA is boosted to compensate for it. This suggests a role of domain-general cognitive resources also in RDA storage, which also casts doubts on the simple notion of the RDA consisting of a small number of discrete slots. On the contrary, information storage in the RDA may be mediated by attentional resources, and thus RDA capacity may be subject to functional variations (Brose et al., 2012).

On the whole, the serial position data suggest that WM contains storage and processing components; that storage components form a storage hierarchy; that storage and processing components may deploy shared cognitive resources; and that components are selectively activated in WM in order to optimize both storage and processing performance. This view is in agreement with the concentric model, but allows for its extension to accommodate general domain cognitive resources and a regulative system for their dynamic allocation.

WM in people with LTM deficits

The structure of WM as suggested by the concentric model (and, specifically, the storage hierarchy) can be effectively studied in people with long-term memory deficits. The APOE-4 genotype is marked by the presence of the allele ε4 of the apolipoprotein E gene. This genotype is the major genetic risk factor for the development of Alzheimer's disease in old age (Raber et al., 2005). There is considerable research on the cognitive profile of healthy APOE-4 carriers which has led to contradictory conclusions. Cognitive performance in young healthy APOE-4 carriers is usually indistinguishable from that in carriers of other genotypes. Healthy APOE-4 carriers may also outperform carriers of other genotypes on some cognitive tasks. However, large-scale meta-analytic studies systematically suggest that APOE-4 carriers have deficits in episodic long-term memory, which is often compensated for by increased cognitive control (Lancaster et al., 2017; Wisdom et al., 2011).

In a study, we assessed WM and its relationship in a sample of healthy APOE-4 carriers and age-matched controls (for details, see Velichkovsky, Roshchina, & Selezneva, 2015) to more fully understand the workings of WM when LTM is not functioning properly. To assess WM, the operation span task (two to six items) and the n-back task (2-back) were administered, along with a battery of cognitive control tasks.

It was found that performance on the operation span and n-back tasks didn't differ between carriers and non-carriers, which suggests typical WM functioning in the carriers. Correlational analyses indicated, however, that the cognitive mechanisms of the WM task performance may differ between the carriers and non-carriers. For instance, it was found that operation spans for different set sizes correlated in the non-carriers, but not in the carriers, indicating a large variety in task execution mechanisms in the latter group. It was also found that operation span and n-back results correlated in the non-carriers but not in the carriers. This again suggests that while in the non-carriers WM tasks execution mechanisms are very consistent, such consistency is absent in the carriers. It was further found that while operation span in the non-carriers correlates with the antisaccade task (an attention control task), such correlations are absent in the carriers.

This is an intriguing result because WM tasks (especially complex span tasks like the operation span task) have been shown to strongly correlate with executive attention tasks like the antisaccade task (Engle, 2002). In this respect it is notable that n-back performance correlated with antisaccade performance in both the carriers and non-carriers. This dissociation may be interpreted in light of the fact that the 2-back task used in this study mostly relies on the RDA for items storage, while storage in the operation span task often exceeds the typical capacity limits of the RDA, and requires the a-LTM for item maintenance. As the a-LTM may be deficient in the carriers, they may use idiosyncratic strategies for the execution of the operation span task, which may thus not exhibit the correlation characteristic of the non-carriers. There are no such idiosyncratic strategies in the execution of the n-back task by the carriers as RDA is intact in them.

These results further corroborate the storage hierarchy view suggested by the concentric model. A fully intact storage hierarchy may be preserved in the non-carriers. This hierarchy includes a short-term memory component (RDA) and a long-term memory component (a-LTM). The latter is used more and more with the progressive increase of the WM load. In the carriers, the a-LTM component is compromised due to a general deficit in episodic LTM in this population. Importantly, the failure of the a-LTM does not preclude the carriers from exhibiting normal performance in WM tasks, especially the operation span task. This may be achieved by the recruitment of RDA resources, better cognitive control, or other specific strategies compensating for the inefficiency of the storage hierarchy. These findings are in strong agreement with the idea of the dynamic resource allocation to the components of WM proposed in the previous section. The functional organization of WM seems to actively compensate for the inefficiencies of selected components optimizing storage and processing during the execution of WM tasks. This again suggests a role for a regulatory component responsible for dynamic resource allocation to be included in the concentric model of WM.

Discussion

Several studies were presented, with the aim of validating the concentric model of WM (Oberauer, 2002). The studies used complex span WM tasks and WM updating tasks, both of which are the gold standard for measuring WM functions (Conway et al., 2005). The studies showed that predictions derived from the concentric model can be largely supported by the data. Below, we consider several of the studies' results and discuss general implications for the theory of WM.

The results imply that WM contains a specialized processing component, which corresponds to the FA identified in the concentric model. Processing in the FA was shown to be independent of information storage in the WM. This follows from the processing complexity factor being systematically statistically independent of the factors affecting storage in the WM. Processing/storage independence may make adaptive sense. Assuming that there are common cognitive resources shared between processing and storage in WM (Towse & Hitch, 1995), the problem arises as to how to prioritize resources in cases where they are functionally or constantly depleted. A priority for processing makes it possible to search for solutions in a dynamic situation even if storage is undermined. Therefore it makes perfect sense to shield processing from the peculiarities of storage implementation. It should be noted that previous correlational studies have also shown a separation of processing from storage in WM (Barrouillet & Camos, 2007; Duff & Logie, 2001; Vergauwe et al., 2014), and suggested that these WM functions are differentially related to cognitive abilities (Unsworth et al., 2009). WM may thus have a modular architecture, with processing and storage modules functioning independently but influencing each other through a set of interfaces (Fodor, 1983).

The results also suggest that there are two functionally different storage systems in WM. These storage systems closely correspond to the distinction between RDA and a-LTM made by the concentric model. One storage system (the RDA) is capable of maintaining only a very limited number of elements (about four items), as evidenced by the absence of interference effects with loads below four items, and a marked decrease in WM recall accuracy for loads over three/four items. This is in strong agreement with research on WM capacity limits, which suggests a "new magic number" of four items (Cowan, 2001). Cowan (2001) reviews an impressive array of research showing that WM capacity was previously overestimated, and that WM capacity converges on about four items in experimental paradigms that preclude mnemonic strategies from better encoding WM content or its transfer to LTM.

On its face, the present result, may be interpreted as indicating storage in the RDA, without the transfer of information to a-LTM (this is precluded by the manipulations described in Cowan, 2001). It is notable that the concentric model also suggests a limit of four items for the RDA (Oberauer, 2002). A related line of research is being pursued by studies of relational complexity (Halford et al., 2005) which suggest that humans can only process relations between four variables. We will consider below why such a stringent capacity limit for RDA may still have adaptive value.

Beside capacity limits, the RDA is marked by its assumed insensitivity to interference effects (Oberauer, 2002). The absence of interference effects in some components of WM is a matter of debate. Some authors argue that parts of WM/short-term memory are immune to interference (Dempster & Cooney, 1982; Halford, Maybery, & Bain, 1988; Tehan & Humphreys, 1995), especially for loads below four items (which relates these findings to the capacity of the FA and the RDA). Other researchers suggest that the FA (in the broader sense including the RDA, Cowan, 1999) is as susceptible to interference as any other memory system (Carrol et al., 2010; Ralph et al., 2011; Shipstead & Engle, 2013).

Our data suggest that the limited-capacity storage system may be immune to interference, as interference effects were observable only outside of this storage system. It should be noted that this result was shown for complex span and updating tasks, which differ from the tasks used to assess the effects of interference, and are arguably more valid measures of WM functions. That items held in the RDA may be shielded from interference may have considerable adaptive value, if one assumes that the RDA's functional role is the storage of several information items most relevant for the solution of the task at hand (in this respect the RDA is akin to the cache memory of modern computer architectures). It would be optimal if the RDA protected its items from decay due to interference or other factors, until they are used in the task-relevant processing, or are explicitly declared to be no-longer relevant and erased from RDA (Ecker et al., 2010; Maxcey & Woodman, 2014). However, it is possible that protection from interference is a dynamic property provided by the inhibitory mechanisms of cognitive control (Engle, 2002). Thus, the presence of interference effects may depend on the availability of inhibition, which may explain the ambivalent results concerning the role of interference in capacity-limited storage.

The second storage system in WM (a-LTM in the concentric model's terms) is less capacity-limited. Given the limitations of our study, we couldn't test for capacity limits of this storage system. The embedded process theory (Cowan, 1999) suggests that there are no such structural limits, and that storage in the a-LTM is limited only by finite activation resources. This view strongly corresponds to the idea that LTM has no storage limits, and that representations can be activated in LTM not only by their activation via the FA, but also due to automatic processes of spreading activation.

The present results further indicate another important feature of the a-LTM; it *is* prone to interference. This makes the a-LTM a typical memory system, and strongly suggests that it is indeed implementedwith the help of LTM mechanisms (see below). Our data also show that the a-LTM is recruited for WM storage only when the capacity limits of the RDA are exceeded by WM storage requirements. This suggests a two-tier storage architecture ("a storage hierarchy"), which is able to satisfy different storage requirements while maintaining quick accessibility of several decay-protected items, and still maintaining access to indefinitely large amounts of information potentially relevant for the solution of the task at hand.

The storage hierarchy view on WM storage which is explicated above, leads to the conclusion that WM storage is partly dependent on LTM storage mechanisms. The involvement of LTM mechanisms in WM has gained some support recently. For instance, research shows clear hippocampal involvement into the execution of WM tasks (the hippocampus being a structure traditionally associated with long-term episodic memory). Faraco et al. (2011) showed hippocampal recruitment during complex span tasks. Hippocampal activation was observed during the operation span task which contrasted to no hippocampal activation during a mental arithmetic task (which is a WM task akin to the processing subtask of the operation span but lacks its storage requirements). Öztekin et al. (2009) suggested a role for the hippocampus during retrieval of items not held in the FA. Leszczynski (2011) reviewed several studies on hippocampal involvement in WM and also suggested that is activated during WM maintenance. The hippocampus can also be involved in WM encoding and updating (Spellman et al., 2015).

Thus, the hippocampus as a LTM-related structure is consistently involved in various aspects of WM, with its involvement in information storage being best documented. This supports behavioral data (for instance, Unsworth et al., 2012) in suggesting that LTM mechanisms are involved in WM. WM is thus not a separate memory system strictly different from LTM. It is a functional system which recruits short-term and long-term memory mechanisms according to the storage/processing requirements as they evolve during the execution of the current task.

Why could the two-tier storage hierarchy be adaptive? First, it can be assumed that RDA capacity limitations are determined by the basic brain architecture (Cowan, 2001)–for instance, by the restrictions imposed by the neurons needed to fire synchronically while maintaining WM content. This leaves only a little capacity for fast-access reliable storage in the RDA, and it is easy to see that this capacity should be used to store the most relevant bits of information. The second storage system would provide access to larger data sets at the cost of slower access and less reliable storage. Thus, the two-tier hierarchy may have grown out of the difficulty of constructing a high-capacity fast-access reliable storage system.

Second, the low capacity of the RDA may be in itself an advantage, if items have to be selected into the FA for further processing (and it seems they have to be selected, Garavan, 1998; Oberauer, 2002). As selection processes are linearly dependent on the size of the set to be selected from, it is advisable to keep the set small if the selection has to be performed quickly. Optimizing selection-for-processing speed is especially important for WM, as it is primarily a system for the support of goal-directed actions. Thus, while increasing RDA capacity may be costly from the brain architecture point of view, it may not even be necessary if selection speed is to be optimized. In this respect the two-tier hierarchy may be an optimal compromise between access speed and information volume. Of course, such an architecture benefits from the putative information exchange mechanisms which enable the transfer of information between the levels of the storage hierarchy (consider the correlation between the a-LTM and the RDA in *Fig. 1*).

Our results on serial position effects further suggest that there may be a reciprocal relationship between the storage and processing components of the concentric model. This may be based on the use of common resources for the execution of specific WM tasks. For instance, the reduction of primacy effects by complex processing which we have observed, indicates that information transition between the RDA and the a-LTM may be resource-consuming. If such resources are diverted to processing in the FA, the off-loading of information items from the RDA to the a-LTM can be prohibited. There is already independent evidence that information transfer within WM storage systems is resource-consuming (Fawcett & Taylor, 2012).

An additionally significant result is the increase of the recency effect which complements the reduction of the primacy effect. This seems to support the view of WM as functional system. Allocation of resources within WM is not static but dynamic; the activation of different WM components is tailored according to the dynamics of task's requirements. Overall, the dynamics of WM functioning and resource allocation optimize storage *and* processing depending on what is more relevant for achieving behavioral goals. This suggests the existence of a control module within WM which regulates the goal-directed resource allocation between WM components. The concentric model lacks such a component, or implicitly embeds it within the FA. The idea of WM as a dynamically regulated functional system is further supported by our data on WM functioning in the APOE-4 carriers, where deficiency in the a-LTM component leads to the functional reorganization of WM tasks performance.

Given that the concentric model gets some empirical support, it is important to relate it to Baddeley's multicomponent WM model, which has for decades dominated WM theory, and for which there is a large amount of evidence. It seems that the concentric model and the multicomponent model are not incompatible but rather complementary.

First, the multicomponent model considers its "slave" (storage) systems to be unitary while the concentric models stress that there are different representational states (Zokaei et al., 2014) of items held in the storage systems. A valuable conclusion from the concentric model is that there is a complex storage hierarchy (or hierarchies) within WM. This conclusion should be incorporated into the multicomponent model.

Second, the multicomponent model stresses that storage in WM is modality-specific. That verbal, visual, and spatial storage processes are distinct within WM is empirically well-supported, but this aspect is lacking in the concentric model and has to be integrated into the concentric model by providing modality-specific storage hierarchies (which opens up an intriguing question of dynamic resource allocation between them, Vergauwe et al., 2012; Velichkovsky & Izmalkova, 2015).

Third, the multicomponent model lacks the FA as a specialized locus of processing, which the concentric model includes. This component should clearly be incorporated into the multicomponent model, given that the FA is clearly independent of storage in WM.

Fourth and last, it has already been shown that the concentric model needs a regulative module for the dynamic allocation of resources, which would enable it to be a functional system dynamically optimized to achieve the correct balance between storage and processing. Such a module is already present in the multicomponent model in the form of the central executive, which plays an important, if underspecified, role. It seems that to be a thorough model of WM, the concentric model will have to explicitly provide for a subsystem akin to the central executive. This would, for instance, explain the close empirical connections between WM capacity and executive control (Engle, 2002; Shipstead, Harrison, & Engle, 2016).

Conclusion

Several studies were reviewed along with results of our recent experiments, with the aim of assessing the validity of the concentric model of WM (Oberauer, 2002; Shipstead, Harrison, & Engle, 2016; Velichkovsky, Nikonova, & Rumyantsev, 2015; Velichkovsky, 2016). The model suggests that WM is comprised of a processing component–the FA, and two functionally distinct storage components–the RDA and the a-LTM. In the present analysis, we demonstrated that the predictions generated from the concentric model were largely supported by the results of several lines of research. In particular, it was found that the FA and the storage system function independently. Furthermore, the results supported the notion of the two storage systems forming a storage hierarchy, with the RDA being a capacity-limited reliable storage of several highly task-relevant items, and the a-LTM being an LTM-based capacity-unlimited storage system providing access to large amounts of potentially task-relevant information. Processing and task complexity in WM were shown to demand common cognitive resources. Overall, human WM proved to be a dynamic system which optimizes the activity of separate components according to the requirements of the current task.

Acknowledgements

This study was supported by the Russian Foundation for Basic Research, grant no. 16-06-00065.

References

Ackerman, P.L., Beier, M.E., & Boyle, M.O. (2005). Working memory and intelligence: The same or different constructs? *Psychological Bulletin, 131*, 30–60. doi: 10.1037/0033-2909.131.1.30

Baddeley, A.D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A.D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29. doi:10.1146/annurev-psych-120710-100422

Barrouillet, P. & Lepine, R. (2005). Working memory and children's use of retrieval to solve addition problems. *Journal of Experimental Child Psychology, 91*, 183–204. doi: 10.1016/j.jecp.2005.03.002

Brose, A., Schmiedek, F., Lövdén, M., & Lindenberger, U. (2012). Daily variability in working memory is coupled with negative affect: the role of attention and motivation. *Emotion, 12*, 605–617. doi: 10.1037/a0024436

Brown, T. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Carroll, L.M., Jalbert, A., Penney, A.M., Neath, I., Surprenant, A.M., & Tehan, G. (2010). Evidence for proactive interference in the focus of attention of working memory. *Canadian Journal of Experimental Psychology, 64*(3), 208–214. doi: 10.1037/a0021011

Case, R., Kurland, M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*, 386–404. doi: 10.1016/0022-0965(82)90054-6

Conway, A., Kane, M., Bunting, M., Hambrick, D., Wilhelm, O., & Engle, R. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786. doi: 10.3758/BF03196772

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory. Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge: Cambridge University Press. doi: 10.1017/CBO9781139174909.006

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–114. doi: 10.1017/S0140525X01003922

Daneman, M., & Carpenter, P.A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*, 450–466. doi: 10.1016/S0022-5371-(80)90312-6

Dempster, F.N. & Cooney, J.B. (1982). Individual differences in digit span, susceptibility to proactive interference, and aptitude/achievement test scores. *Intelligence, 6,* 399–416. doi: 10.1016/0160-2896(82)90026-5

Ecker, U.K.H., Lewandowsky, S., Oberauer, K, & Chee, A.E.H. (2010). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology, 36,* 170–189. doi: 10.1037/a0017891

Engle, R. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science, 11,* 19–23. doi: 10.1111/1467-8721.00160

Faraco, C.C., Unsworth, N., Langley, J., Terry, D., Li, K., Zhang, D., … & Miller, L. S. (2011). Complex span tasks and hippocampal recruitment during working memory. *Neuroimage, 55*(2), 773-87. doi: 10.1016/j.neuroimage.2010.12.033

Fawcett, J. M. & Taylor, T.L. (2012). The control of working memory resources in intentional forgetting: Evidence from incidental probe word recognition. *Acta Psychologica, 139*, 84–90. doi: 10.1016/j.actpsy.2011.10.001

Fodor, J.A. (1983). *Modularity of mind: An essay on faculty psychology*. Cambridge, Mass.: MIT Press.

Fox, J. (2006). Structural equation modeling with the sem package in R. *Structural Equation Modeling, 13,* 465–486. doi: 10.1207/s15328007sem1303_7

Garavan, H. (1998). Serial attention within working memory. *Memory & Cognition, 26*, 263–276. doi: 10.3758/BF03201138

Gathercole, S.E. & Alloway, T.P. (2008). *Working memory and learning: A practical guide*. Thousand Oaks: Sage Press.

Giofre, D., Mammarella, M., & Cornoldi, C. (2013). The structure of working memory and how it relates to intelligence in children. *Intelligence, 41,* 396–406. doi: 10.1016/j.intell.2013.06.006

Hale, S., Rose, N., Myerson, J. Strube, M., Sommers, M., Tye-Murray, N., & Spehar, B. (2011). The structure of working memory abilities across the adult life span. *Psychology and Aging, 26,* 92–110. doi: 10.1037/a0021483

Halford, G.S., Baker, R., McCredden, J.E., & Bain, J.D. (2005). How many variables can humans process? *Psychological Science, 16,* 70–76. doi: 10.1111/j.0956-7976.2005.00782.x

Halford, G.S., Maybery, M.T., & Bain, J.D. (1988). Set-size effects in primary memory: An age-related capacity limitation? *Memory & Cognition, 16,* 480–487. doi: 10.3758/BF03214229

Kane, M., Conway, A., Hambrick, D., & Engle, R.W. (2007). Variation in working-memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake & J. Towse (Eds.), *Variation in working memory* (pp. 21–48). New York: Oxford University Press.

Lancaster, C., Tabet, N., & Rusted, J. (2017). The elusive nature of APOE ε4 in mid-adulthood: Understanding the cognitive profile. *Journal of the International Neuropsychological Society, 23,* 239–253. doi: 10.1017/S1355617716000990

Lemaire, P. & Fayol, M. (1995). When plausibility judgments supersede fact retrievals: The example of odd-even effect on product verification. *Memory & Cognition, 23*(1), 34–48. doi: 10.3758/BF03210555

Lepine, R., Bernardin, S., & Barroillet, P. (2005). Attention switching and working memory spans. *European Journal of Cognitive Psychology, 17*(3), 329-345. doi: 10.1080/09541440440000014

Leszczynski, M. (2011). How does hippocampus contribute to working memory processing? *Frontiers in Human Neuroscience, 5,* 168. doi: 10.3389/fnhum.2011.00168

Maxcey, A. & Woodman, G. (2014). Can we throw information out of visual working memory and does this leave informational residue in long-term memory? *Frontiers in Psychology, 5,* 294. doi: 10.3389/fpsyg.2014.00294

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(3), 411–421. doi: 10.1037/0278-7393.28.3.411

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin Review, 19*(5), 779–819. doi: 10.3758/s13423-012-0272-4

Öztekin, I., McElree, B., Staresina, B.P., & Davachi, L. (2009). Working memory retrieval: Contributions of the left prefrontal cortex, the left posterior parietal cortex, and the hippocampus. *Journal of Cognitive Neuroscience, 21*, 581–593. doi: 10.1162/jocn.2008.21016

Owen, A.M., McMillan, K.M., Laird, A.R., & Bullmore, E. (2005). N-Back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping, 25,* 46–59. doi: 10.1002/hbm.20131

Raber, J., Huang, Y., & Ashford, J.W. (2005). ApoE genotype accounts for the vast majority of AD risk and AD pathology. *Neurobiology of Aging, 25*, 641–650. doi: 10.1016/j.neurobiolaging.2003.12.023

Ralph, A., Walters, J.N., Stevens, A., Fitzgerald, K.J., Tehan, G., Surprenant, A.M., … & Turcotte, J. (2011). Immunity to proactive interference is not a property of the focus of attention in working memory. *Memory & Cognition, 39*(2), 217–30. doi: 10.3758/s13421-010-0030-7

Schweppe, J., Grice, M., & Rummer, R. (2011). What models of verbal working memory can learn from phonological theory: Decomposing the phonological similarity effect. *Journal of Memory and Language, 64,* 256–269. doi: 10.1016/j.jml.2010.11.006

Shipstead, Z. & Engle, R.W. (2013). Interference within the focus of attention: Working memory tasks reflect more than temporary maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(1), 277–289. doi: 10.1037/a0028467

Shipstead, Z., Harrison, T.L., & Engle, R.W. (2016). Working memory capacity and fluid intelligence: Maintenance and disengagement. *Perspectives on Psychological Science*, *11*(6), 771–799. doi: 10.1177/1745691616650647

Spellman, T., Rigotti, M., Ahmari, S.E., Fusi, S., Gogos, J.A., & Gordon, J.A. (2015). Hippocampal-prefrontal input supports spatial encoding in working memory. *Nature, 522*(7556), 309–14. doi: 10.1038/nature14445

Sternberg, S. (1984). Stage models of mental processing and the additive-factor method. *Behavioral and Brain Sciences. 7,* 82–84. doi: 10.1017/S0140525X00026285

Tehan, G. & Humphreys, M.S. (1995). Transient phonological codes and immunity to proactive interference. *Memory & Cognition, 23*(2), 181–191. doi: 10.3758/BF03197220

Towse, J. & Hitch, G. (1995). Is there a relationship between task demand and storage space in tests of working memory capacity? *Quarterly Journal of Experimental Psychology, 48*, 108–124. doi: 10.1080/14640749508401379

Turner, M. & Engle, R. (1989). Is working memory capacity task dependent? *Journal of Memory & Language, 28*, 127–154. doi: 10.1016/0749-596X(89)90040-5

Unsworth, N., Redick, T., Heitz, R., Broadway, J., & Engle, R. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between processing and storage. *Memory, 17,* 635–654. doi: 10.1080/09658210902998047

Unsworth, N., Spillers, G., & Brewer, G. (2012): Working memory capacity and retrieval limitations from long-term memory: An examination of differences in accessibility. *The Quarterly Journal of Experimental Psychology, 65*(12), 2397–2410. doi: 10.1080/17470218.2012.690438

Velichkovsky, B.B. (2016). Short- and long-term storage components in working memory. *International Journal of Psychology, 51*(S1), 180.

Velichkovsky, B.B. & Izmalkova, A.I. (2015). Vlijanie nagruzki na verbalnuju rabochuju pamjat pri glazodvigatelnoj aktivnosti v uslovijah vypolnenija zadanija zritelnogo poiska [Effect of verbal working memory load during the oculomotor activity in visual search]. *Eksperimentalnaya psikhologiya [Experimental Psychology], 8*(2), 21–35.

Velichkovsky, B.B., Nikonova, E.Yu., & Rumyantsev, M.A. (2015). Struktura rabochej pamjati: effekty uvelichenija kolichestva informacii, uslozhnenija obrabotki i interferencii [The structure of working memory: Effects of increased information load, processing complexity, and interference]. *Psikhologicheskii zhurnal [Psychological Journal], 36*(3), 38–48.

Velichkovsky, B.B., Roschina, I.F., & Selezneva, N.D. (2015). Cognitive control and memory in healthy apoe-ε4 carriers with a family history of Alzheimer's disease. *Psychology in Russia: State of the Art, 8*(1), 4–13. doi: 10.11621/pir.2015.0101

Vergauwe, E., Dewaele, N., Langerock, N., & Barrouillet, P. (2012). Evidence for a central pool of general resources in working memory. *Journal of Cognitive Psychology, 24*, 359–366. doi: 10.1080/20445911.2011.640625

Wisdom, N.M., Callahan, J.L., & Hawkins, K.A. (2011). The effects of apolipoprotein E on non-impaired cognitive functioning: A meta-analysis. *Neurobiology of Aging, 32*, 63–74. doi: 10.1016/j.neurobiolaging.2009.02.003

Zokaei, N., Ning, S., Manohar, S., Feredoes, E., & Husain, M. (2014). Flexibility of representational states in working memory. *Frontiers in Human Neuroscience, 8,* 853. doi: 10.3389/fnhum.2014.00853

Original manuscript received March 19, 2016
Revised manuscript accepted August 29, 2016
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society
Lomonosov Moscow State University

Discourse abilities in the structure of intelligence

Anatoly N. Voronin[a], Olga M. Kochkina[b*]

[a] *Institute of Psychology, Russian Academy of Sciences, Moscow, Russia*
[b] *Moscow State Univeristy of International Affairs, Moscow, Russia*

* Corresponding author. E-mail: olkochkina@yandex.ru

Background. This article is devoted to empirical research on discourse abilities within the structure of cognitive abilities. Discourse abilities, as well as linguistic abilities, are part of language abilities, but they are directly linked with discourse practices and a certain communicative situation. Discourse abilities allow a person to effectively initiate, keep, expand, and complete the process of communication, using language appropriate to any given situation. These abilities contribute to making communication more effective and achieving mutual understanding between partners, while at the same time they speed up the process of forming an interaction strategy. The empirical verification of the construct "discourse abilities," and the design of original diagnostic tests on them, led us to differentiate linguistic and discourse abilities.

Objective. However, it is not yet clear what place discourse abilities occupy in the structure of cognitive abilities. This is the primary goal of our research.

Method. The design of the study involved group testing (in groups of 15-35 people) using the following methods: a discourse abilities test; a short selection test; a social intelligence test, and short variations of Torrance's and Mednick's tests. In total, 208 people (133 women and 75 men, ages 17 to 21 years) participated in the study, all of them either first year humanities students or high school students from Moscow.

Results and Discussion. The research results revealed that discourse abilities relevantly correlate with the majority of indicators of general and social intelligence and creativity (except non-verbal intelligence). Discourse abilities as part of the structure of cognitive capabilities form a discrete factor, and include relevant components such as verbal and general intelligence and indicators of social intelligence, such as the ability to group expressions. Structures indicative of cognitive abilities varied within the study group, which included people with different levels of discourse abilities. A data structure which conformed to an a priori structure of cognitive abilities was observed only in the group with the medium level of discourse abilities. The group with a low level of discourse abilities mostly showed the aggregation of various indicators of intelligence and creativity, while the group with a high level of discourse abilities showed further differentiation of intelligence types, and the evolution of discourse abilities into a separate factor.

Key words: human cognition, structure of intelligence, psychometrics, creativity, discourse, discourse abilities, communication

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0306
http://psychologyinrussia.com

Introduction

The ability to master language is the vital part of the human mind. Modern societies are estimated to have increasingly engaged verbal intelligence more than non-verbal. Language ability is a constellation of psychological and physiological conditions which ensures the understanding, and adequate reproduction of language signs by the members of the language community (Leontev, 2014). Language abilities determine the ease with which linguistic knowledge and rules of analysis and synthesis of language units are acquired, which allows constructing and analyzing sentences, and using the language system for communication purposes. These abilities contribute to the speed with which a language (both native and non-native) is mastered, and also to the effectiveness of language use in communication (Kabardov, 2003).

General language abilities are usually subdivided into two components: the linguistic one, which provides for the mastery of the language base, and the communicative one, responsible for successful communication. The former is more essential, as it implies the mastery of language unit models, rules of word changes and collocations, and general vocabulary; the latter, being linked with not strictly linguistic phenomena such as pronunciation, word alteration, and collocation variants, choice of synonyms, etc., is not as binding, and individual peculiarities are possible (Smirnitskii,1981). Factors such as emotional expressiveness, motivation, and the speakers' intentions are often ignored although they define the individuality and communicative aspects of speech, and, consequently, should be part of those studies of discourse that deal with the text immersed in communication (Arutyunova, 1998). Any discourse is simultaneously directed toward the situation in which it occurs (the socio-cultural context sets the rules for conducting conversation and its forms of expression), and toward the person being spoken with (interlocutors communicate, influence each other, and express their opinions, intentions, and views concerning the situation). Discourse is a form of single, partner- and milieu-coordinated verbal behavior, supported by a complex knowledge system, depending on the communicative competence of the speakers (Pavlova, 2002). Discourse abilities thus can be defined as abilities to master and realize discursive practices, which are carried out on two levels: as a mental representation of the current social situation, and a representation of a collective subject, obtained through cultural and historical experience (Voronin & Kochkina, 2008). These abilities allow for the enhanced effectiveness of interaction and more adequate mutual understanding among people in the process of communication; besides, they accelerate the process of defining a strategy for cooperation.

The notion of "discourse abilities" is closely linked with the notion of "communicative competence." Communicative competence is a person's ability to adequately arrange his/her speech in productive and receptive ways, with the help of language usage corresponding to any concrete situation (Zimnyaya, 1989), as well as to combine social, national, and cultural modes, assessments, and values which determine not only a suitable form, but also acceptable content (Vereshchagina & Kostomarov, 1982).Thus, discourse abilities can be viewed as the operationalized part of communicative competence, which allows the initiation, support, expansion, and conclusion of the process of communication with the help of situationally appropriate verbal means.

In the course of empirically verifying the "discourse abilities" construct, and developing linguistic and discourse ability diagnostic tests, we succeeded in differentiating linguistic and discourse abilities (Voronin & Kochkina, 2009). At the same time the discourse abilities scale, designed on the basis of the English language, has pronounced limitations: it is an English version of how to diagnose discourse abilities exercised while studying English as a foreign language. The Russian language method for diagnosing discourse abilities was created at the end of 2013, and is based on data about the modern city communication (Kitaigorodskaya & Rozanova, 2003). Our research suggests that there are different types of modern communicative space: communication at home and outdoors; purpose-oriented communication and factual communication; weekday communication (holiday communication, working-time communication); and free-time communication. Each communicative area is characterized by a specific form of discourse.

In defining the types of discourse for the developing material of this study, we proceeded from the types of modern city communicative space mentioned above, and took into consideration the psychological peculiarities of everyday discourse (Zachesova & Grebenshchikova, 2007). The material of the method reflects nine types of discourse: 1) humorous announcements; 2) phone conversations; 3) family discourse; 4) business discourse (mostly between employer and employee); 5) teacher-student situational discourse; 6) internet discourse; 7) discourse used in talking with strangers; 8) communicating with people in the service sector; and 9) discourse with friends. Validation of the method revealed that only four generalized types of discourse are verified and valid: discourse when communicating with strangers and acquaintances; discourse when cooperating with relatives and friends; business discourse and internet discourse. These are the types of discourse through which discourse abilities were revealed in our study.

It is possible to define subject areas closely related to the concept of "discourse abilities." The notions that are the closest semantically to "discourse abilities" are the following: "general intelligence," or the successful functioning of the person as a whole; "verbal intelligence," or the ability to carry out verbal mental analysis and synthesis to solve verbal tasks, define notions, determine similarities, etc. (i.e. ability to master the language); and "social intelligence," or the ability to cognize social phenomena (Kochkina, 2009). Actually, the goal of our work is to determine the position of discourse abilities within the structure of such cognitive capabilities.

Method

Research design and procedure

Our study of the combined structure of intelligence, creativity, and discourse abilities was carried out in 2013-2014, and involved first-year humanities students from the GAUGN (State Academic University for Humanities) and the Moscow Institute of Economics, Politics and Law, and high school students from Moscow school 539 and gymnasium 1503. The overall number of the participants in the study was 208 (133 women and 75 men, ages 17 to 21). The study involved tests in groups of 15 to 35 students using the following methods: a Discourse abilities test; a short selection test adapted by Buzin (Buzin, 1989); a social intelligence test from G. Gilford and

M. Sullivan adapted by Mikhailova (Mikhailova, 2006); and short variations of Torrance's and Mednick's tests adapted by Voronin and Galkina (Voronin & Galkina, 1994; Voronin, 1994).

Methods

The level of intelligence of the participants was estimated with the help of the short selection test adapted by V N. Buzin for fast diagnosis of the following abilities: the ability to summarize and analyze material; flexibility of thinking; inertia in thinking and the ability to change subjects; emotional components of thought and distractibility, speed and accuracy of perception, distribution and concentration of attention; language usage and grammatical correctness; choice of optimum strategy, and spatial imagination (Buzin, 1989).

The study's design demanded that the intellectual productivity evaluation be carried out on three indicators (verbal intelligence, non-verbal intelligence, and an integral indicator) by grouping test points according to the types of the stimulus material. The level of social intelligence was estimated with the help of the adapted version of the G. Guilford and M. Sullivan test adapted by Mikhailova (Mikhailova, 2006). Verbal and non-verbal creativity was diagnosed through short variations of Torrance's and Mednick's tests (Voronin & Galkina, 1994; Voronin, 1994). Creativity was estimated by several indicators: productivity, originality, uniqueness, and flexibility. The method of diagnosing discourse abilities was based on everyday vocabulary (Voronin, 2014). The afore-mentioned method suggests that the study subjects read a short description of a certain communicative situation and choose the answer most closely corresponding to the described situation. Below we present some examples.

2. You disliked the latest book you read so much that you wrote on the Internet forum: "Disgusting work, this can hardly be called literature." You get the response: "Speculations of an immature person, nothing more can be added."

Your answer is …

1) Can't catch up with you wise old farts;
2) Every person has his own opinion, so I don't consider mine wrong;
3) That's it–a mature man will never read this garbage;
4) Immature people are illiterate, and this book is really bad, and you, sir, have clearly failed to read it;
5) Try to convince me to change my opinion; I might not have noticed its value;
6) I'm describing my feelings. And I can't like all the books in the world.

5. You answer the phone at home and you hear, "My dear friend, could I please speak to Mr. Ivanov?" You understand that your father is being asked for. You call him saying,

1) Comrade, you are wanted on the phone!
2) Da-a-a-ad!
3) You find out who it is and tell your father, "Dad, Mr. Petrov is calling."
4) One moment…;
5) Mr. Ivanov, this is for you;
6) Dad, it's for you.

The participant's choice allowed us to draw conclusions about his/her ability to initiate, keep up, develop, and conclude communication, using language that is appropriate to the situation, — i.e., discourse abilities.

Results and Discussion

Our results were processed with the help of SPSS Statistics 19.0. In the course of the processing, descriptive statistics were estimated, and correlations and factor analyses of the data were carried out. The distribution of variations in discourse abilities turned out to be pseudo-normal, with two additional peaks in the area of low and high values respectively. Discourse abilities turned out to be closely linked with almost all indicators of intelligence and creativity (Table 1).

Table 1. Correlations τ (tau) Kendall of discourse ability indicators with different types of intelligence and creativity

Indicators	Discourse abilities
Social intelligence. Cartoon Predictions	.31**
Social intelligence. Expression grouping	.45**
Social intelligence. Verbal expression	.19**
Social intelligence. Missing Cartoons	.29**
Social intelligence. Composite score	.34**
SST[1] verbal IQ	.55**
SSTnon-verbal IQ	0.04
SSTtotal score IQ	.22**
SSTpoints reviewed	.23**
Creativity by Mednick. Productivity	.23**
Creativity by Mednick. Originality	.15**
Creativity by Torrance. Fluency	.16**
Creativity by Torrance. Originality	.17**
Creativity by Torrance. Elaboration	.21**
Creativity by Torrance. Flexibility	.11*

** — Correlation significant at $p < .01$; * — Correlation significant at $p < .05$

The discourse abilities test is a method of revealing verbal peculiarities in the cognitive sphere; therefore its correlation with "verbal" indicators of intelligence and creativity is expected to be higher than that with "non-verbal" indicators. This statement seemed to be true with respect to intelligence. Correlation between the indicators of discourse abilities and verbal intelligence is nearly significant at a level

1 SST — short selection test.

of p<.01, while correlation with non-verbal intelligence is negligible. The highest correlation between the indicators is between verbal intellect and discourse ability. Analysis of social intelligence correlations shows the reverse tendency.

At the same time all the correlations between indicators of discourse abilities and social intelligence indicators are positive and relevant. Correlations with creativity indicators are also positive and significant. The sole indicator that is not linked with discourse abilities is non-verbal intelligence. Analysis of the factor structure of discourse abilities, intelligence, and creativity factors indicates (Table 2) that five major factors can be distinguished: 1) social intelligence, 2) non-verbal creativity, 3) general intelligence, 4) verbal creativity, and 5) discourse abilities. These are most tightly connected with verbal intelligence, general intelligence, and one of the factors of social intelligence — the ability to "expression grouping".

Table 2. Matrix of factor solutions for indicators of intelligence, creativity, and discourse abilities

	Component				
	1	**2**	**3**	**4**	**5**
Discourse abilities	.274	.254	–.011	.137	.755
Social IQ Cartoon Predictions	.863	.096	.104	.058	.189
Social IQ Expression grouping	.624	.113	.024	.027	.538
Social IQ Verbal expression	.485	–.172	.243	.455	.199
Social IQ Missing Cartoons	.822	.077	.099	.127	.087
Social IQ Composite score	.911	.034	.148	.111	.279
SST verbal IQ	.208	.038	.391	–.015	.776
SST non-verbal IQ	.178	.097	.903	.022	–.155
SST IQ	.247	.090	.866	.007	.347
SST points reviewed	–.113	.189	.684	–.091	.432
Cr Mednick productivity	.122	.098	–.252	.714	.129
Cr Mednick originality	–.037	.071	.183	.785	.019
Cr Torrance fluency	.251	.888	–.038	.134	.094
Cr Torrance originality	.099	.740	.291	.313	.104
Cr Torrance elaboration	.684	.378	–.051	–.180	–.124
Cr Torrance flexibility	.018	.900	.136	.042	.157

The factor structure which was revealed in the analysis basically reproduced the cognitive ability structure, which includes indicators of discourse abilities according to the method based on the English language (Kochkina, 2009). That study revealed four factors: the general intelligence factor; the general intelligence and linguistic ability factor; the social intelligence factor; and the discourse abilities and social intelligence factor. Thus discourse abilities proved to be significantly linked with general intelligence, verbal intelligence, and social intelligence ("verbal expression" scale).

The high correlation between discourse abilities and major indicators of different types of intelligence and creativity could be interpreted as bringing out some basic, primary capability, which describes the intellectual sphere as a whole, and comprises both reproductive and productive intelligence. The most adequate and suitable answers of the participants in various communicative situations may be a sign of the kind of rational (intellectual) scheme which they use in daily situations, where rational and reasonable verbal reactions are uncommon. This interpretation of discourse abilities is in line with the concept of practical intelligence identified in R.J. Sternberg's theory of "intellect leading to success" (Sternberg, 2002; Sternberg, Kaufman, 1998). The latter states that practical intelligence predetermines the realization of ideas, and ensures success in a certain social group. It is also expressed in the manifold structure of the factor of discourse ability, which includes indicators of general and verbal intelligence and the factor of social intelligence, shown by "expression grouping".

Consequently, discourse abilities can be psychometrically interpreted as a verbal manifestation of intelligence used to evaluate another person's state, feelings and emotions, and intentions. In other words, it describes the cognitive faculty, which is better known in the literature as "theory of mind". Additionally a person with a high level of discourse abilities can verbalize non-verbal communication components more effectively. Our interpretation of discourse abilities based on the empirical data we obtained suggests that a high level of discourse ability defines the following characteristics of a person:

- Completeness, accuracy, and flexibility in describing a stranger's personality;
- Sensitivity to other people's emotional states in business communication;
- Variety of expressiveness in communication;
- Openness and friendly disposition in communication;
- Sensitivity to feedback in communication, receptivity to criticism;
- High self-esteem;
- Varied and complex description of self-image;
- Exact understanding of how one's own emotional state is perceived by one's communicative partners, which indicates congruence of communicative behavior;
- Adequate situations of self-presentation.

For the present moment the afore-mentioned characteristics are mostly speculations and need further empirical verification. Another, more plausible interpretation of the results may be linked to special features of the sample and the vagueness of the instructions used in testing discourse ability. It did not seem possible to conduct post factum an additional empirical study aimed at revealing the particulars of the subjects' comprehension of the instructions, but additional analysis of the structure of the data obtained could be done. For this purpose the sample was divided into three groups according to the level of discourse abilities they expressed: a group with high indicators (upper quartile), a group with low indicators (lower quartile), and the rest — people who were tested with a medium level of discourse ability. Afterwards we did correlational and factor analyses of these groupings.

Correlations between the level of discourse ability, and intelligence and creativity indicators in the different groups, are shown in Table 3.

Table 3. Correlations τ (tau) Kendall of discourse ability indicators with different types of intelligence and creativity

	Discourse abilities		
	Low level	**Medium level**	**High level**
Social IQ Cartoon Predictions	.578*	.357*	–0.057
Social IQ Expression grouping	.734*	.380*	0.051
Social IQ Verbal expression	.306*	.403*	0.092
Social IQ Missing Cartoons	.359*	.359*	–0.14
Social IQ Composite score	.595*	.452*	–0.034
SST verbal IQ	.325*	.395*	–0.159
SST non-verbal IQ	0.08	–0.113	–0.287
SST IQ	0.226	0.139	–.301*
SST points reviewed	–0.009	0.057	–0.119
Cr Mednick productivity	–0.217	0.091	0.1
Cr Mednick originality	.353*	.204*	0.022
Cr Mednick uniqueness	0.129	0.089	.242*
Cr Torrance fluency	.656*	–.214*	–0.069
Cr Torrance originality	.647*	–.339*	.618*
Cr Torrance elaboration	.400*	.212*	–0.191
Cr Torrance flexibility	.589*	–.412*	–.272*

* — Correlation significant at $p < .05$

Correlation analysis shows that the group with a high level of discourse abilities has a different structure of correlation interdependencies than the others. A negative correlation between general intelligence and discourse abilities was revealed: there were no connections with social intelligence, but there were significant connections with the major indicators of verbal and non-verbal creativity (with originality and uniqueness), with various indicators of creativity being linked with discourse abilities in different ways. A similar diversity of creativity indicators' ties in various testing situations has also been revealed previously–for instance, when the interconnection of creativity indicators and intelligence of teachers and pupils was studied (Voronin, 2004; Voronin & Trifonova, 2003).The correlations of various indicators in groups with medium and low level of discourse abilities are more comparable. There were positive interrelations between discourse abilities, and social intelligence and verbal creativity indicators. Correlations with general intelligence are non-significant. There are pronounced differences in the case of non-verbal creativity indicators: the group with a low level of discourse ability shows a significant positive correlation, and the group with medium level shows a negative one.

All three groups' data underwent factor analysis by the method of principal components with the consequent Rotation Method: Varimax with Keiser Normalization. Factor analysis results for the group with low discourse abilities are given in Table 4.

Table 4. Matrix of factor solutions for indicators of intelligence and creativity for the sample showing low-level discourse abilities

	Component				
	1	**2**	**3**	**4**	**5**
Discourse abilities	.414	.701		.122	–.183
Social IQ Cartoon Predictions	.890	.321	–.116		
Social IQ Expression grouping	.533	.672		–.212	–.340
Social IQ Verbal expression	.813	.196	.180	–.289	.287
Social IQ Missing Cartoons	.895	–.105	.211		
Social IQ Composite score	.938	.313		–.107	
SST verbal IQ		.569	.279	–.366	.460
SST non-verbal IQ	.253		.883		–.138
SST IQ	.266	.193	.906	–.102	.100
SST points revised	–.145	.168	.884		.201
Cr Mednick productivity		–.189			.920
Cr Mednick originality		.271		.844	–.178
Cr Mednick uniqueness		.239		.861	.387
Cr Torrance fluency	.261	.849		.257	–,197
Cr Torrance originality		.844	.424	.228	
Cr Torrance elaboration	.669	.269	.304	.124	–.229
Cr Torrance flexibility		.841	.122	.299	.121

Factor analysis of the results revealed five factors, whereby only two of them could be easily interpreted: those of social intelligence and creativity. General intelligence proved to be divided into two factors: the general and non-verbal intelligence factor, and the verbal intelligence factor, which is linked with verbal creativity. The latter is comprised of non-verbal creativity, verbal intelligence, and one of the indicators of social intelligence, the ability to expression grouping. It is the factor on which discourse abilities have the highest loading. At the same time discourse abilities are closely linked with other indicators of social intelligence and verbal creativity. The contribution of discourse abilities to the verbal creativity factor is negative.

Factor analysis of the sample showing a medium level of discourse abilities (Table 5) clearly indicates four factors: 1) social intelligence, 2) general intelligence, and 3) non-verbal and 4) verbal creativity. Factor five is comprised of social intelligence and verbal creativity readiness. Discourse abilities with high positive loadings are included in both factors of social intelligence, and their contribution to non-verbal creativity factor is negative.

Table 5. Matrix of factor solutions for indicators of intelligence and creativity for the sample showing medium-level discourse ability

	Component				
	1	2	3	4	5
Discourse abilities	.564		–.512		.272
Social IQ Cartoon Predictions	.237	.394		.120	.800
Social IQ Expression grouping	.733	.114		.115	.218
Social IQ Verbal expression	.767	.200	–.107	.235	
Social IQ Missing Cartoons	.647		.150		.542
Social IQ Composite score	.632	.241			.643
SST verbal IQ	.429	.584	–.204	–.182	.135
SST non-verbal IQ		.882	.147		
SST IQ	.187	.955			.120
SST points reviewed	.377	.568	.211	–.443	–.237
Cr Mednick productivity				.866	.205
Cr Mednick originality	.177			.792	–.204
Cr Mednick uniqueness	.237		.285	.889	
Cr Torrance fluency		–.234	.805	.204	.369
Cr Torrance originality		.240	.799	.276	.104
Cr Torrance elaboration			.171	–.123	.822
Cr Torrance flexibility			.924		

The results of the high discourse ability sample (Table 6) indicate that discourse abilities constitute a separate factor, which is closely linked with the major indicator of non-verbal creativity — originality. Besides that, five more factors may be distinguished: 1) social intelligence, linked with the non-verbal creativity readiness indicator; 2) verbal creativity, with a high loading of the social intelligence indicator of verbal expression; 3) general and non-verbal intelligence, with a negative loading of indicators of verbal creativity elaboration; 4) general and verbal intelligence, with a loading of social intelligence indicators (grouping of expressions); and 5) non-verbal intelligence. Discourse abilities have negative loading for two factors: those of non-verbal creativity, and general and non-verbal intelligence.

Table 6. Matrix of factor solutions for indicators of intelligence and creativity for the sample showing high-level discourse abilities

	Component					
	1	**2**	**3**	**4**	**5**	**6**
Discourse abilities		–.219		–.157		.889
Social IQ Cartoon Predictions	.953		.166		.125	
Social IQ Expression grouping	.746		–.261	–.110	.468	.187
Social IQ Verbal expression	.305		.774		.354	.157
Social IQ Missing Cartoons	.778	.284		.176	–.241	–.185
Social IQ Composite score	.953		.181		.198	
SST verbal IQ		.286		.199	.783	–.112
SST non-verbal IQ		.949		.129		
SST IQ	.122	.860		.207	.417	
SST points reviewed	.517	.513		.298	.514	
Cr Mednick productivity		–.645	.592		–.339	–.130
Cr Mednick originality		.462	.775	–.162	.265	.140
Cr Mednick uniqueness			.922	.153	–.250	.149
Cr Torrance fluency		.271		.830	.268	.246
Cr Torrance originality	–.115	.273	.303	.262		.836
Cr Torrance elaboration	.798	–.151	–.367		–.346	–.157
Cr Torrance flexibility				.965		–.150

Factor analysis conducted in the groups with different levels of discourse abilities revealed that interconnections between discourse abilities and various indicators of intelligence and creativity in the different groups vary. The factor structure of the data for different groups also varies. These contradictory results can be accounted for in the following way.

First, as a person's intelligence level increases, differentiation within his/her cognitive abilities are enhanced (Voronin, 2004; Druzhinin, 2001). In groups with low and medium discourse abilities, fewer factors are distinguished, and indicators of various types of intelligence and creativity are interlinked to a greater degree. A data structure conforming to an a priori structure of cognitive capabilities is noticeable only in the group with a medium level of discourse abilities. Integration of various indicators of intelligence and creativity is mostly present in groups with a low level of abilities, and groups with a high level of capabilityies reveal further differentiation of intellect types, with discourse ability becoming a separate factor.

Secondly, the data concerning discourse abilities can be explained by the vagueness of the task to be solved while doing the discourse ability test: the instructions required the choice of an answer most closely corresponding to the given situation. Under conditions of such vagueness, examinees with a low level of discourse abilities might have been trying to solve the given task in a creative way based on their social intelligence. Medium-level examinees might have also based their answers on their social intelligence, but followed non-verbal stereotypes concerning the given situations. The groups with high levels of discourse ability treated the task given by the instructions literally (to choose the answer that corresponds to the situation most precisely), and solved it in a creative way based on their image of the situation.

The diversity of the structure of cognitive capabilities which depend on the level of discourse abilities lets us speak of the specific characteristics of the verbal behavior exhibited by people whose discourse abilities differ. The verbal communication of a person with a low level of discourse abilities will be spontaneous, based on guesses concerning causes of interaction, without a clear view of the partner's aims and motives. Such people's communication is determined by their social experience. The spontaneity of events will direct communication of people with a low level of discourse ability.

People with a medium level of discourse abilities base their communication on a stereotypical view of various communicative situations. The scope of past experience in dealing with communicative situations will account for the success of verbal interaction. Hence communication problems arising from that approach stem from the absence of a holistic vision of the situation, and the impossibility of changing the situation throughout the conversation. Predetermination is an attribute of discourse of people with a medium level of discourse abilities.

When discourse abilities are highly developed, they become one of the most significant factors determining verbal behavior. It might be stated that verbal behavior adequate to the situation appears at this very stage. At the same time such success stems from a situational view of the ongoing communication, and implies creative development of verbal interaction: there might be a change of initiative in the dialogue, and changes in the tempo, content, and purport of the talk in accordance with the situation and the participants' intentions. The possibility of varying the discourse is due to the high level of cognitive capabilit~~y~~ies, and the creative intentions of the interlocutors.

Conclusion

Discourse abilities significantly correlate with the majority of the indicators of general and social intelligence and creativity (except non-verbal intelligence). Discourse abilities as part of the structure of cognitive ability forms a discrete factor with such relevant components as verbal and general intelligence, and indicators of social intelligence, such as the ability to group expressions. Structural indicators of cognitive capabilities vary in samples with different levels of discourse ability. A data structure which conforms to an a priori structure of cognitive capabilities is observed only in the group with a medium level of discourse abilities. The group with a low level of discourse abilities mostly shows aggregation of various indicators of intelligence and creativity, while the group with a high level of discourse abilities shows further differentiation of intelligence types, and the evolution of discourse abilities into a separate factor.

References

Arutyunova, N.D. (1998). Diskurs. In *Yazykoznanie. Bolshoi entsiklopedicheskii slovar* [Linguistics. Big encyclopedian dictionary] (p. 137). Moscow: Bolshaya Rossiiskaya Entsiklopediya.

Buzin, V.N. (1989). Kratkii orientirovochnyi test [Short selection test]. In *Praktikum po psikhodiagnostike. Konkretnye psikhodiagnosticheskie metodiki* [Applied handbook on psychological assessment. Psychodiagnostic methods]. (pp.112–126). Moscow: Moscow University Press.

Druzhinin, V.N. (2001). *Kognitivnye sposobnosti: struktura, diagnostika, razvitie* [Cognitive abilities: structure, diagnostics, development]. Moscow.

Kabardov, M.K. (2003). *Kommunikativnye i kognitivnye sostavlyayushchie yazykovykh sposobnostei (individualno-tipologicheskii podkhod)* [Communicative and cognitive components of language abilities (individual-typological approach)]. (Doctoral dissertation). Psychological Institute of the Russian Academy of Education, Moscow.

Kitaigorodskaya, M.V., & Rozanova, N.N. (2003). Sovremennoe gorodskoe obshchenie: Tipy kommunikativnykh situatsii I ikh zhanrovaya realizatsiya (na primere Moskvy) [Modern urban communication: types of communicative situations and their genre realization (on the example of Moscow)]. (pp. 103–120). *Sovremennyi russkii yazyk: Sotsialnaya I funktsionalnaya organizatsiya.* Moscow: Yazyki Slavyanskoi Kultury.

Kochkina, O. M. (2009). *Struktura i dinamika yazykovykh sposobnostei v diskursivnoi praktike* [Structure and dynamics of language abilities in discourse practice]. (PhD dissertation). State Academic University of Humanities, Moscow.

Leontev, A.A. (2014). *Slovo v rechevoi deyatelnosti* [Word in speech activity]. Moscow: URSS. (Original work published 1965)

Mikhailova, E.S. (2006). *Test J. Guilforda i M. Sullivana. Diagnostika sotsialnogo intellekta.Metodicheskoe rukovodstvo* [Guilford and Sullivan's test. Methodical guide]. Saint Petersburg: IMATON.

Pavlova, N.D. (2002). Kommunikativnaya paradigma v psikhologii rechi i psikholingvistike [Communicative paradigm in speech psychology and psycholinguistics]. In N.D. Pavlova 9Ed.), *Psikhologicheskie issledovaniya diskursa* [Psychological research of discourse]. (pp. 7–17). Moscow: PerSe.

Smirnitskii, A.I. (1981). Objektivnost sushchestvovaniya yazyka [Objectivity of language existence]. *Istoriya sovetskogo yazykoznaniya. Nekotorye aspekty obshchei teorii yazyka: Khrestomatiya* [The history of Soviet linguistics. Some aspects of the general theory of language: A reader]. (pp. 64–72). Moscow: Vysshaia Shkola.

Vereshchagina, E.M., & Kostomarov, V.G. (1983). *Yazyk I kultura: Lingvostranovedenie v prepodavanii russkogo yazyka kak inostrannogo* [Language and culture: Linguistics in teaching Russian as a foreign language]. (p. 257). Moscow: Russkii Yazyk.

Voronin, A.N. (2014). *Metodika diagnostiki diskursivnykh sposobnostei na material povsednevnoi leksiki* [Methods of diagnosing discourse abilities basing on everyday vocabulary material]. *Eksperimentalnaya psikhologiya [Experimental Psychology], 7*(2), 94–112.

Voronin, A.N., & Galkina, T.V. (1994). *Diagnostika verbalnoi kreativnosti (adaptatsiya testa Mednika)* [Diagnostics of verbal creativity (adaptation of Mednick's test)]. In A.N. Voronin

(Ed.), *Metody psikhologicheskoi diagnostiki* [Methods of psychological diagnostics]. (pp. 40–81). Moscow: Institute of Psychology, Russian Academy of Sciences.

Voronin, A.N. (2004). *Intellekt i kreativnost v sovmestnoi deyatelnosti* [Intelligence and creativity in joint cooperation] (Doctoral dissertation). Institute of Psychology, Russian Academy of Sciences, Moscow.

Voronin, A.N. (1994). *Diagnostika neverbalnoi kreativnosti (kratkii variant testa Torrensa)* [Diagnostics of nonverbal creativity (abridged version of Torrance's test)]. In In A.N. Voronin (Ed.), *Metody psikhologicheskoi diagnostiki* [Methods of psychological diagnostics]. (pp. 5–40). Moscow: Institute of Psychology, Russian Academy of Sciences.

Voronin, A.N., & Kochkina, O.M. (2008). *Diskursivnye i lingvisticheskie sposobnosti v structure intellekta cheloveka* [Discourse and linguistic abilities in human intelligence structure]. *Psikhologiya. Zhurnal Vysshei Shkoly Ekonomiki [Psychology. Journal of the Higher School of Economics], 2*, 124–132.

Voronin, A.N., & Kochkina, O.M. (2009). *Struktura yazykovykh sposobnostei v protsesse osvoeniya diskursivnoi praktiki* [Discourse abilities structure in the process of mastering discourse practice]. *Eksperimentalnaya psikhologiya [Experimental Psychology], 2*(3), 92–110.

Voronin, A.N. & Trifonova, I.G. (2003). *Sootnoshenie skhodstva/razlichiya lichnostnykh osobennostei uchitelya I uchenika s kreativnostyu uchashchikhsya* [Correlation of similarity/differences of personal peculiarities of teacher and pupil with pupils' creativity]. *Psikhologicheskiy zhurnal [Psychological Journal], 24(6),* 77–85.

Zachesova, I.A., & Grebenshchikova, T.A. (2007). *Vzaimootnosheniya sobesednikov v dialoge* [Interlocutors' relationships in a dialogue]. In N.D. Pavlova & I.A. Zachesova (Eds.), *Situatsionnaya i lichnostnaya determinatsiya diskursa* [Situational and personal determination of discourse]. Moscow: Institute of Psychology, Russian Academy of Sciences.

Zimnyaya, I.A. (1989). *Psikhologiya obucheniya nerodnomu yazyku* [Psychology of teaching a foreign language]. Moscow: Russkii Yazyk.

Original manuscript received November 14, 2015
Revised manuscript accepted June 27, 2016
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society · Lomonosov Moscow State University

The unconscious nature of insight: A dual-task paradigm investigation

Anton A. Lebed[a,b]*, Sergei Y. Korovkin[a,c]

[a] *Yaroslavl State University, Yaroslavl, Russia*

[b] *University of Delaware, Newark, DE, USA*

[c] *Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia*

* Corresponding author. E-mail: coglebed@gmail.com

Background. Insight is a specific part of the thinking process during creative problem solving. The experience of a sudden unexpected solution of the problem makes it distinct from other problem solving. Though the insight problem solving process is hidden from the observer and the solver himself, it is possible to study working memory changes during the problem-solving process in order to observe the tracks of insight.

Objective. A critical experiment was carried out to determine whether it is legitimate to measure insight-problem-solving dynamics within a dual-task paradigm and working memory model. Also a verification was conducted of the hypothesis of whether insight problem solving competes for cognitive resources with unconscious processes.

Design. We designed a special procedure based on Kahneman's (1973) modified dual-task paradigm, allowing simultaneous performance of the problem-solving process and probe tasks of different types. The reaction time was measured for the probe task. There were two problems conditions (insight and regular), and two probe tasks conditions (implicit and explicit). Participants: 32 participants, aged from 18 to 32 years ($M = 19.81$; $\sigma = 2.51$).

Results. Significant differences in implicit probe reaction time were found between the dual-task condition (implicit categorization and insight problem solving) and solo implicit probe condition ($t(15) = -3.21$, $p = .006$, $d = -.76$). A joint effect of problem type and probe type was found ($F(1, 60) = 4.85$, $p = .035$, $\eta_p^2 = .07$).

Conclusion. The results support the idea that information processing of conscious and of unconscious processes are separate. Unconscious processing capacity is limited. Implicit skill seems to be operated by the same mechanisms as insight problem solving, therefore competing for a common resource. It was also shown that such hidden creative unconscious processes as insight can be tracked via working memory load.

***Keywords*:** insight, thinking, dual-task, implicit learning, working memory, problem solving

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0307
http://psychologyinrussia.com

Introduction

Insight has always been one of the most mysterious phenomena in the psychology of thinking. Its mechanisms, role, and structure are still uncertain. Some investigators consider that there is no special problem type known as "insight". The cybernetic model of Newell and Simon (1972) attempted to describe the role and processing of insightful solutions in the terms of regular problem solving. They suggested that there is no specific method of insight problem solving, so it can be explained in terms including operators, heuristics, and problem space. However, many critics of this view have noticed that creative problem solving is often indescribable in the common terms of a cybernetic model, as it lacks visible structure.

Many other investigators have approached insight as a special specific phenomenon in a separate class of creative problems known as insight problem solving. Researchers vary a lot, for example using different stimuli, but methods that are completely alike. Chu and MacGregor (2012) present an overview of the best-known methods and problems used to investigate insight. They made a catalog of problems, varying in difficulty and representation type: classical problems such as the six matchstick problem, verbal riddles, and spatial puzzles. These problems vary greatly, making it hard to integrate data from several experiments into one conclusion. The new problems include matchstick arithmetic and compound remote associations. The authors compare theories of insight, concluding that there is no complete answer yet to the problem of insight.

Since the identification of insight as a subject for psychological study, there have been few investigations of its mechanisms and structure. The main reason for this lies in its unconscious nature, making it very difficult to gather data about it. Even the problem-solver cannot realize how far he is from the solution, as shown in the "feeling of warmth" investigation by Metcalfe and Wiebe (1987). They showed that participants were unable to adequately measure how close they are to the answer in insight problem solving, even though they gave rather precise estimates during regular problem solving.

One of the key questions in understanding the nature of insight consists in figuring out what processes precede insight, what makes it more difficult, and what allows a person to overcome an impasse in the problem-solving process. Empirically, insight problem solving is described in terms of multiple impasses that the solver encounters after using all available options, followed by incubation, which is characterized by very limited conscious thinking. There are various points of view on the nature of incubation and concurring processes.

On the one hand, there are doubts about the very existence of some specific active processes in incubation. According to this point of view, incubation itself consists of nothing special, nothing more than forgetting incorrect solution attempts and options (Anderson, 2010; Simon, 1977; Woodworth, 1938). This approach assumes that incubation is a temporary period, during which heuristics, solving strategies, and after-effects of incorrect solutions are erased. Mental fixation, occurring in early stages of problem solving, can be overcome by just forgetting actual schemas; therefore, the effectiveness of problem solving is directly related to the duration of incubation (Dodds, Ward, & Smith, 2004; Sio & Ormerod, 2009). Some believe that incubation is required for an attention switch, rather than forgetting

(Segal, 2004). The attention switch is supposed to be relatively fast and spontaneous. This idea predicts the absence of a positive relation between problem-solving effectiveness and incubation duration. In memory erase models, working memory is considered to play only a minor role. There are other possible explanations for the lack of visible activities during incubation within the "nothing special" approach: unsolved problems might be stored in memory in order to look for problem-relevant environmental cues that would make it possible to solve the problem using new information (Seifert, Meyer, Davidson, Patalano, & Yaniv, 1995). According to this point of view, incubation is considered an active process of awaiting relevant information that requires some usage of long-term working memory.

On the other hand, there are models of insight problem solving incubation that assume specific unconscious processes that are distinct from conscious mechanisms. There are multiple data suggesting that a solution can be found before the solver is aware of it. Thus, anagram investigation with eye-tracking showed that participants tend to focus their attention on the solution part of the anagram some time prior to the actual solution (Ellis, 2012; Ellis, Glaholt, & Reingold, 2011). These authors conclude that there is unconscious knowledge of the solution prior to the solution itself. Additionally, there is evidence that unconscious hints can greatly affect insight problem solving (Bowden, 1997; Thomas & Lleras, 2009; Werner & Raab, 2013), while verbal hints are ineffective (Weisberg & Alba, 1981). The time at which the hint is presented can also affect its effectiveness. The solver has to spend some time working on the problem to be able to understand the hint, while spending too much time increases the effect of incorrect solutions (Moss, Kotovsky, & Cagan, 2011). When solving arithmetical problems, children are capable of using new methods and solving strategies long before they become aware of this ability and are able to report it (Siegler, 2000).

Even considering substantial evidence supporting the idea of unconscious processes of insight incubation, there is no common model of unconscious information processing in problem solving. One approach is based on the idea of two separate parallel, yet competing, systems: System 1 and System 2 (Kahneman, 2011) or conscious and unconscious modes of thought (Dijksterhuis & Nordgren, 2006). According to these models, drawing attention from or overloading conscious thought can facilitate unconscious information processing. At the same time, conscious (logical) and unconscious (intuitive) thinking can be considered not only as independent systems, but also as levels/layers of one process (Ponomarev, 1976). During the creative problem-solving process, the solver "climbs" from infantile forms of intellect towards more ontologically mature methods. In case of failure to solve a problem, the solver returns to the lowest levels of thinking. Another approach attempts to describe the processes underlying incubation in terms of semantic networks and neural networks (Hélie & Sun, 2010; Martindale, 1995; Sio & Rudowicz, 2007). The process of incubation is considered a gradual or rapid change of activation in the network and the creation of distant associations. This approach explains the effect of semantic hints during problem incubation, while other effects are much less readily interpretable in these terms. The mechanisms of unconscious processing are usually described in general terms and are hard to prove experimentally. Some of the known mechanisms of insight solution, such as constraint relaxation (disabling rules, supported by functional fixedness) and chunk decomposition (dividing up pieces of information that are perceived as a whole) (Knoblich, Ohlsson, & Raney, 2001; Öllinger, Jones, Faber, & Knoblich, 2013) are hard to consider either exclusively conscious or unconscious. One of the potential reasons why the data acquired by different researchers are so controversial is the use of different stimulus materials (problems of different sorts), different incubation models, and different approaches to affect incubation. Additionally, different stages of problem solving might have different underlying processes. These issues lead to the question, whether it is possible to investigate thinking processes using one single problem scenario.

Korovkin, Vladimirov, and Savinova (2014) investigated the insight-problem-solving process within a dual-task paradigm, showing the differences between insight problem solving and regular problem solving. The participants were asked to solve either an insight or regular (algorithmized) problem and to perform a probe task at the same time. The solution time was divided into 10 stages, each representing the average reaction time for a probe task at different moments of the solution. They found that there is a significant difference between the last stages of insight problem solving and regular problem solving.

Other experiments investigated different mental storage systems by varying the probe task material and the problem types. Some experiments were designed to discover whether the information contained in working memory blocks can be put into a state of competition by attaching probe tasks of various kinds.

However, some have criticized the procedures of such investigations. According to cognitive unconscious theory (Allakhverdov, 2009), there are two big obstacles to measuring insight within a dual-task paradigm. First of all, unconscious processes are considered limitless by cognitive unconscious theory, making it impossible to create any competition between insight problem solving and the probe task. Secondly, the author pointed to the conscious nature of working memory, concluding that it is not legitimate to investigate an unconscious process — insight — by having it compete with working memory probe tasks.

Indeed, the original working memory models considered this memory as fully conscious. Working memory was introduced by Baddeley and Hitch (1974) as a system providing temporary storage for manipulations of information required for complex cognitive problem solving. It consists of central executive control and three subsystems: a phonological loop that processes verbal data; a visuospatial sketchpad that processes visual data and spatial relations; and an episodic buffer that combines all the types of information to create a working space. Later on, several attempts were made to investigate the role of working memory in unconscious thinking. Some authors used an implicit learning paradigm to study this. Being an unconscious process, implicit learning might give investigators a clue about the role of working memory load in the learning process (Reber, 1967). It was discovered that working memory overload negatively affects implicit learning (Reber, & Kotovsky, 1997) and that implicit skill might be stored in working memory (Hassin, Bargh, Engell, & McCulloch, 2009).

In our studies, we use Kahneman's (1973) resource competition model as well as a probe task method. The resource model supposes that there is a limited cognitive resource, used by most mental processes that compete for it. According to its author, this resource has a biological basis: arousal. In a situation of resource deficit, task evaluation occurs, so the most relevant tasks receive the most resources. He demonstrated his model with a simple experiment using a dual-task paradigm: simultaneous processing of mental problem solving and a simple probe task. It turned out that the effectiveness of solving the probe task dramatically decreases in the middle of the primary problem solving, because this period requires the greatest amount of resources. Based on the amount of mistakes (in the probe task) in this period, he concluded that all available resources were allocated to the primary problem, leaving the probe task unsupported.

Method

According to the above-mentioned investigations, we decided that implicit learning might be a fitting probe task to interact with the insight-problem-solving process through the means of working memory load. The experimental procedure combines two paradigms: the working memory competition paradigm and the dual-task paradigm, allowing us to investigate the thinking process during the problem-solving process based on reaction time and mistakes in the probe task performed at the same time as the participant solves the problem.

There are several requirements for any probe task. It should be:

1) *Rapid*. The more often the participant is required to react for the probe task, the more descriptive dynamics we can receive.
2) *Congruent*. The probe task must be relevant to the experimental problem in order to achieve some interaction with it.
3) *Simple*. The probe task should not distract from the problem-solving process; otherwise, the whole solution will be uninformative, because of its unnatural processing.

Our main goal was to find out whether there is competition between unconscious processes and insight problem solving. We have chosen implicit skill to represent unconscious processes.

According to the dual-task paradigm, the participants had to perform the problem solving and the probe task at the same time. As the participant reads and thinks about the primary problem, the probe task appears on the screen, requiring constant reactions for changing stimuli. The probe task is usually a simple decision or categorization task that requires minimum thinking and limited input variations of two or three buttons. Based on the reaction time of the probe task, we can judge participants' mental tension (representing working memory load) at the different stages of the solution process. The greater the reaction time for the probe task, the less free resources are available for it because of the increased resource-requirements of the primary task. This method allows us to observe the dynamics of the thinking process in detail, but in this paper, we discuss the efficiency of the probe task performance without respect to its dynamics.

The study sample consisted of 32 participants, aged from 18 to 32 years ($M = 19.81$; $\sigma = 2.51$), with 7 men and 25 women: students, graduate students, and other people with higher education. Participants received no additional motivation. They were initially seated approximately 45 cm from the monitor, but were free to move their heads during the experiment; therefore, the visual angle subtended by the probe task was not controlled. Stimulus size varied from 5 cm to 9 cm depending on the categorization features.

To test the validity of the dual-task paradigm application in insight problem solving, we hypothesized that the probe task can compete with insight problem solving for common resources. Based on the idea of two relatively separate systems of information processing in the human mind, we also hypothesized that insight problem solving uses specific unconscious/implicit working memory resources.

To verify this hypothesis, we created a computer program in the Python environment with PsychoPy application v1.76.00 (Peirce, 2007). We varied the working memory load type with the probe task: either implicit with unconscious selection criteria or explicit with conscious selection criteria. We also varied the problem type: either insight or regular problems. The problems can be found in Appendix A. We measured probe task reaction time and average solution time. There were four experimental groups: implicit probe task with insight problem, explicit probe task with regular problem, implicit probe task with regular problem, and explicit probe task with insight problem.

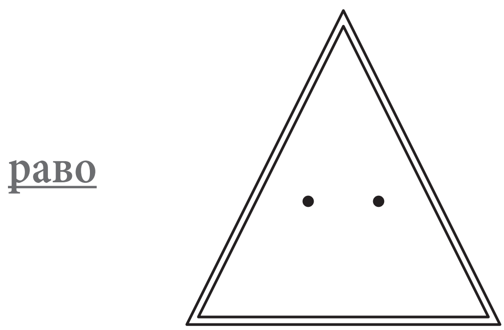


Figure 1. Probe tasks examples (verbal probe task and figures probe task)

The probe task examples are illustrated in Figure 1. There were two kinds of probe tasks, to avoid stimulus-specific results: figures and nonsense syllables. The participants were asked to group both figures and nonsense syllables into either "left" or "right" categories, based on a certain rule; the probe stimulus could vary in color, form, size, and additional markers. Only a certain combination of these properties was considered *right*, the rest were considered *left*. The categorization criteria are stated in Appendix B.

The rules were shown to the participants in the "explicit" groups and were hidden from those in the "implicit" groups. During a practice trial, the implicit group performed an implicit learning sequence. The participants were asked to categorize the stimuli and had audial feedback: a ring-sound if the answer was correct and a drum-sound if the answer was wrong. The sequence was considered complete when the participants were able to produce correct answers in no less then 59% of the trials. Participants passed two trials in our experiment: the first one was a practice (control) trial and the second one was experimental. During the practice trial, the participants had to become familiarized with the experimental stimuli or to develop an implicit skill (in the implicit probe task group). Once the participants were familiar with the probe task, the second trial began: simultaneous performing the probe task and the primary problem. The participants had to push buttons to categorize the stimuli that popped up and to solve the primary problem.

There were either creative insight problems, or regular problems that required simple mathematical calculation. Regular problems feature distinctive sequential steps that are required to solve them. These steps are known to the solver, along with all possible actions within the problem space. The procedure/sequence for the solution can be described in terms of algorithms. Insight problems, on the contrary, require a substantial change of problem representation, possible actions or their applications (a functional solution is required), while the final goal, representation, or possible actions are not initially obvious to the solver. An example of an insight problem: "Misha and Sasha played in the basement. It was dark and dirty in there. Once they got upstairs, Sasha's face turned out to be very dirty, but Misha's face was clean. Nevertheless, only Misha went to wash his face, but not Sasha. Why?" An example of a regular problem: "Three chickens lay three eggs in three days. How many eggs will 12 chickens lay in 12 days?" There were two problems of each type, to avoid problem-specific results. All the problems had been chosen as having the same success rate and solution time. The problems were presented in the form of a text on the screen.

Results

Statistical analysis was performed using two-way ANOVA and t-test for dependent and independent samples. We received significant results, providing evidence on the researched goal: there is a significant difference or probe task reaction time between the control condition of implicit practice and the experimental condition of simultaneous performance of implicit probe task with insight problem ($t(15) = -3.21$, $p = .006$, $d = -.76$). Average reaction time for a probe task is significantly higher in dual-task conditions compared to control single-task conditions (see Table 1). This means that insight problem solving and the implicit probe task compete for the cognitive resources of working memory. But surprisingly, we no found differences between practice conditions and the experimental condition of simultaneous performance of implicit ($t(15) = -.27$,

Table 1. Average reaction time of probe task performance

	Implicit probe task		**Explicit probe task**	
	Insight problem	**Regular problem**	**Insight problem**	**Regular problem**
Conditions	M (SD)	M (SD)	M (SD)	M (SD)
Practice (without problem solving)	1.07(0.44)	1.24 (0.49)	2.64 (0.95)	2.52(0.95)
Experiment(with problem solving)	1.55(0.78)	1.3(0.79)	2.37(1.11)	1.99 (0.88)

Note. M — mean, SD — standard deviation

$p = .79$, $d = -.06$) and explicit ($t(15) = .93$, $p = .37$, $d = .26$) probe tasks with regular problems. Counterintuitive results were found in comparison between explicit practice and the experimental condition of simultaneous performance of an explicit probe task with an insight problem ($t(15) = 2.94$, $p = .01$, $d = .58$). In the case of an explicit probe task in insight problem solving, we found that the reaction time in the dual-task condition is significantly less. We propose that the probes were too difficult to perform them simultaneously with primary problem solving. Thus, our participants might have stopped paying attention to the difficult secondary probe task altogether.

We further examined how problem type and probe task type affect the average solution time using ANOVA (see Table 2). We found that probe task type significantly affects the average solution time ($F(1, 60) = 4.85$, $p = .035$, $\eta_p^2 = .07$)[1]. At the same time, the problem type effect is not significant ($F(1, 60) = .16$, $p = .69$, $\eta_p^2 = .002$) and there is no joint effect of both probe task type and problem type ($F(1, 60) = 2.83$, $p = .098$, $\eta_p^2 = .05$).

Table 2. Average solution time of regular and insight problems

	Insight problem	Regular problem
Probe task	**M (SD)**	**M (SD)**
Implicit	354.14 (276.94)	169.19 (178.16)
Explicit	400.28 (242.52)	513.92 (580.52)

Note. M — mean, SD — standard deviation

Pairwise comparison using *t*-test shows significant differences in the average solution time between regular problem solving during implicit probe task performance and other conditions: regular problem with explicit probe task ($t(30) = -2.27$, $p = .03$, $d = -.8$), insight problem with explicit probe task ($t(30) = -3.07$, $p = .005$, $d = -1.09$), and insight problem with implicit probe task ($t(30) = -2.25$, $p = .03$, $d = -.79$). No other significant differences between groups were found.

Conclusion

As we can see from the results by comparing the practice and dual-task conditions, implicit learning probe performance was significantly impaired by the presence of insight problem solving. This kind of interaction of similar information type can be explained by resource competition in working memory. Though the original working memory is considered to have only three blocks for situational information — a phonological loop, a visuospatial sketchpad, and an episodic buffer — they do not account for consciousness of this information. As there is no division into conscious and unconscious information types in working memory model, it cannot fully explain the competition we found in our experiment. However there are extensions to the classical working memory model: Global Workspace theory

[1] η_p^2 — Partial eta-squared (effect size).

and Intelligent Distributive Agent theory (Baars, & Franklin, 2003). Global Workspace theory suggests that there is a special workspace where conscious perception, imagery, inner speech, and reportable goals are processed and kept. Other working memory blocks are used in "consciousness cycles", sending unconscious information into the working space. In terms of an updated working memory model, an interaction of conscious and unconscious processing is possible by means of competition for resource (Intelligent Distributive Agent), which is supported by our empirical data. Another working memory model that accounts for implicit processes features the Implicit Working Memory construct (Hassin et al. 2009). Implicit Working Memory was shown to be involved in performing tasks outside of awareness. It seems that insight problem solving involves a great deal of implicit processing (as shown by competition with implicit learning), leading to the conclusion that implicit working memory tasks can be successfully used as a probe in a dual-task paradigm studying insight problem solving.

On the contrary, there was no difference between explicit probe performance in training conditions and in dual-task conditions. This might have two reasons: assuming that decision making using explicit rules is primarily a System 2 activity, it features, firstly, fast learning (which allows compensation for dual-task conditions by rapid increase in skill), and, secondly, conscious control availability (which allows participants to maintain the same performance by decreasing precision in the probe task). Both suggestions, however, require further empirical verification.

Other results suggest that participants had substantial difficulties with insight problem solving while performing implicit categorization, and with regular problem solving while performing explicit categorization. Regular and insight problem solving seem to have different involvement in System 1 and System 2. Because these two systems (referred to as Default Network and Control Network) were shown to be very distinctive in terms of activation (Gu et al., 2015), two dual-task activities that are processed within the same system (e.g., insight problem solving and implicit learning) have to share and compete for activation, whereas performing activities processed by different systems (e.g., implicit learning and regular problem solving) involves much less interference. In other words, conscious and unconscious information processing have distinct capacities and underlying neurological bases. Conceptually, regular problem solving might rely more on such conscious features, as attention or working memory, while insight problem solving requires implicit and/or bottom-up processing. This suggestion is consistent with the common view of the unconscious and sudden nature of insight.

Further experiments should aim at a more careful and precise load of working memory blocks in order to obtain data about insight solution requirements in working memory.

Limitations

The limitations of our study are common for working memory model investigations and insight problem solving. Working memory is a complex system that is hard to affect precisely, loading one of its blocks and avoiding others. Insight problem solving is often accompanied by discussions of where the problem stimuli could really be called insightful, as there are no complete descriptions and common views on insight problems. The issue of the problem stimuli can be solved either by using only problems that are commonly considered insightful or by introducing a list of criteria for such problems. The other issue of insight problem is the lack of a "eureka" experience in some participants, although they successfully solved the insight problem.

Future directions

There are multiple possible prospective directions of this study; first, having an implicit dual-task probe as a tool to interact with unconscious information processing, the dynamic features of the insight incubation processes can be investigated and discovered. Second, dual-process theories can greatly benefit from this method, as it allows dissociating System 1 (implicit processing) from System 2 (explicit processing) in a broad variety of activities.

The results obtained lead to the conclusion that the probe task method is appropriate for investigation of insight problem solving by means of creating working memory competition between a probe task and problem solving leading to measurable change in reaction time. Therefore, it can be specifically used to investigate the problem of unconscious incubation. The probe task, however, should have a low complexity level to avoid withdrawing too much of the resource (working memory) from the primary task and thereby interfering with the thinking process. Developing such probe tasks will allow us to observe a stable, visible working memory load pattern during the insight incubation process and, therefore, is a goal of our future research. Distinctive features of the probe task should be relevant to the primary task — an insight problem — and, more specifically, to the processes that occur during its incubation.

Acknowledgements

This work is supported by the Russian Science Foundation (16-18-10030) and the State Assignment (25.5666.2017/BP).

References

Allakhverdov, V.M. (2009). *Razmishleniya o nauke psihologii s vosklicatel'-nim znakom.* [Discussions about psychology]. Saint Petersburg: Format.

Anderson, J.R. (2010). *Cognitive Psychology and its implications* (7th ed.). New York: Worth Publishers.

Baars, B.J., & Franklin, S. (2003). How conscious experience and working memory interact. *Trends in Cognitive Science*, 7(4), 166–172. doi: 10.1016/S1364-6613(03)00056-1

Baddeley, A.D., & Hitch, G.J. (1974).Working memory. In G.H. Bower (Ed.). *The Psychology of Learning and Motivation*. Vol. 8. NY: Academic Press, 47–89. doi: 10.1016/S0079-7421-(08)60452-1

Bowden, E.M. (1997). The effect of reportable and unreportable hints on anagram solution and the Aha! experience. *Consciousness and Cognition*, 6(6), 545–573. doi:10.1006/ccog.1997.0325

Chu, Y., & MacGregor, J.N. (2011) Human performance on insight problem solving: A review. *The Journal of Problem Solving,* 3(2), 119–150. doi: 10.7771/1932-6246.1094

Dijksterhuis, A., & Nordgren, L.F. (2006). A theory of unconscious thought. *Perspectives on Psychological Science,* 1(2), 95–109. doi:10.1111/j.1745-6916.2006.00007.x

Dodds, R.A., Ward, T.B., & Smith, S.M. (2004). A review of the experimental literature of incubation in problem solving and creativity. In M.A. Runco (Ed.), *Creativity research handbook*. (3rd ed., pp. 251–284). Cresskill, NJ: Hampton.

Ellis, J.J., Glaholt, M.G., & Reingold, E.M. (2011). Eye movements reveal solution knowledge prior to insight. *Consciousness and Cognition,* 20(3), 768–776. doi: 10.1016/j.concog.2010.12.007

Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q.K., Alfred, B.Y., Kahn, A.E., ... & Bassett, D.S. (2015). Controllability of structural brain networks. *Nature Communications*, 6. doi:10.1038/ncomms9414

Hassin, R.R., Bargh, J.A., Engell, A.D., & McCulloch, K.C. (2009). Implicit working memory. *Consciousness and Cognition*, 18(3), 665–678. doi: 10.1016/j.concog.2009.04.003

Hélie, S., & Sun, R. (2010). Incubation, insight, and creative problem solving: A unified theory and a connectionist model. *Psychological Review*, 117(3), 994–1024. doi:10.1037/a0019532

Kahneman D. (1973) *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Knoblich, G., Ohlsson, S., Raney, G.E. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, 29(7), 1000–1009. doi: 10.3758/BF03195762

Korovkin, S., Vladimirov, I., & Savinova, A. (2014). The dynamics of working memory load in insight problem solving. *The Russian Journal of Cognitive Science*, 1(4), 67–81.

Martindale, C. (1995). Creativity and connectionism. In S. M. Smith, T.B. Ward, & R.A. Finke (Eds.), *The Creative Cognition Approach* (249–268). Cambridge, MA: MIT Press.

Metcalfe, J., Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition,* 15(3), 238–246. doi: 10.3758/BF03197722

Moss, J., Kotovsky, K., & Cagan, J. (2011). The effect of incidental hints when problems are suspended before, during, or after an impasse. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 37(1), 140–148. doi:10.1037/a0021206

Newell, A., Simon, H.A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Öllinger, M., Jones, G., Faber, A.H., & Knoblich, G. (2013). Cognitive mechanisms of insight: the role of heuristics and representational change in solving the eight-coin problem. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 39(3), 931–939. doi: 10.1037/a0029194

Peirce, J.W. (2007) PsychoPy — psychophysics software in Python. *Journal of Neuroscience Methods,* 162(1–2), 8–13. doi: 10.1016/j.jneumeth.2006.11.017

Ponomarev, Y.A. (1976). *Psikhologiya tvorchestva* [The Psychology of creativity]. Moscow, Nauka.

Reber, A.S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior,* 6(6), 855–863. doi: 10.1016/S0022-5371(67)80149-X

Reber, P.J., & Kotovsky, K. (1997). Implicit learning in problem solving: The role of working memory capacity. *Journal of Experimental. Psychology: General.* 126(2). 178–203. doi: 10.1037/0096-3445.126.2.178

Segal, E. (2004). Incubation in insight problem solving. *Creativity Research Journal*, 16(1), 141–148. doi:10.1207/s15326934crj1601_13

Seifert, C.M., Meyer, D.E., Davidson, N., Patalano, A.L., & Yaniv, I. (1995). Demystification of cognitive insight: opportunistic assimilation and the prepared-mind perspective. In R.J. Sternberg & J.E. Davidson (Eds.), *The Nature of Insight* (65–124). Cambridge, MA: MIT Press.

Siegler, R.S. (2000). Unconscious insights. *Current Directions in Psychological Science*, 9(3), 79–83. doi:10.1111/1467-8721.00065

Simon, H.A. (1977). Scientific discovery and the psychology of problem solving. In *Models of Discovery and other topics in the methods of science* (pp. 286–303). Springer Netherlands. doi:10.1007/978-94-010-9521-1_16

Sio, U.N., & Ormerod, T.C. (2009). Does incubation enhance problem solving? A meta-analytic review. *Psychological Bulletin*, 135(1), 94–120. doi:10.1037/a0014212

Sio, U.N., & Rudowicz, E. (2007). The role of an incubation period in creative problem solving. *Creativity Research Journal*, 19(2–3), 307–318. doi:10.1080/10400410701397453

Thomas, L.E., & Lleras, A. (2009). Covert shifts of attention function as an implicit aid to insight. *Cognition*, 111(2), 168–174. doi:10.1016/j.cognition.2009.01.005

Weisberg, R.W., & Alba, J.W. (1981). An examination of the alleged role of "fixation" in the solution of several "insight" problems. *Journal of Experimental Psychology: General,* 110(2), 169–192. doi:10.1037/0096-3445.110.2.169

Werner, K., & Raab, M. (2013). Moving to solution: effects of movement priming on problem solving. *Experimental Psychology*, 60(6), 403–409. doi:10.1027/1618-3169/a000213

Woodworth, R.S. (1938). *Experimental psychology.* New York: Holt.

Original manuscript received May 15, 2015
Revised manuscript accepted June 06, 2016
First published online September 30, 2017

Appendix A. Problems

Regular problems:

1) Calculate: 25*65 = ?
2) Three chickens produce three eggs in three days. How many eggs would 12 chicken produce in 12 days?

Insight problems:

1) Misha and Sasha were playing in the attic. It was dark and dirty in there. When they came downstairs, Sasha's face was covered with dust, while Misha's face was clean. Nevertheless, only Misha decided to wash his face. Why?
2) A magician has put 11 coins on the table. He asks spectators to remove five coins out of 11 and add four coins in such way that nine coins remain. How should they do this?

Appendix B. Rules of categorization in explicit probe task

For both categorization stimuli, there was a single rule to determine the correct category of the presented object. Each object had four features with three possible options within each feature. Only one of three options was considered correct, except for one feature that had two possible correct options to decrease the learning difficulty. If the presented object was correct according to two or more features, it was considered correct for the categorization task.

Specifically, the features and their correct options (in bold) were the following:

For words stimuli

1) Word length — **3 letter, 4 letters**, 5 letters
2) Size — small, **medium**, huge
3) Font — italic, bold, **regular**
4) Color — **green**, black, red

For figures stimuli

1) Shape — **circle, square**, triangle
2) Border lines — **one thin line**, two thin lines, one thick line
3) Dots inside — no dots, **one dot**, two dots
4) Color — **black**, green, red

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society
Lomonosov Moscow State University

COGNITON IN TECHNOLOGICAL CONTEXT

Gaze-and-brain-controlled interfaces for human-computer and human-robot interaction

Sergei L. Shishkin[a], Darisii G. Zhao[a,b],
Andrei V. Isachenko[a,b], Boris M. Velichkovsky[a,b,c*]

[a] *Kurchatov Institute National Research Center, Moscow, Russia*

[b] *Moscow Institute for Physics and Technology, Moscow, Russia*

[c] *Russian State University for the Humanities, Moscow, Russia*

* Corresponding author. E-mail: velich@applied-cognition.org

Background. Human-machine interaction technology has greatly evolved during the last decades, but manual and speech modalities remain single output channels with their typical constraints imposed by the motor system's information transfer limits. Will brain-computer interfaces (BCIs) and gaze-based control be able to convey human commands or even intentions to machines in the near future? We provide an overview of basic approaches in this new area of applied cognitive research.

Objective. We test the hypothesis that the use of communication paradigms and a combination of eye tracking with unobtrusive forms of registering brain activity can improve human-machine interaction.

Methods and Results. Three groups of ongoing experiments at the Kurchatov Institute are reported. First, we discuss the communicative nature of human-robot interaction, and approaches to building a more efficient technology. Specifically, "communicative" patterns of interaction can be based on joint attention paradigms from developmental psychology, including a mutual "eye-to-eye" exchange of looks between human and robot. Further, we provide an example of "eye mouse" superiority over the computer mouse, here in emulating the task of selecting a moving robot from a swarm. Finally, we demonstrate a passive, noninvasive BCI that uses EEG correlates of expectation. This may become an important filter to separate intentional gaze dwells from non-intentional ones.

Conclusion. The current noninvasive BCIs are not well suited for human-robot interaction, and their performance, when they are employed by healthy users, is critically dependent on the impact of the gaze on selection of spatial locations. The new approaches discussed show a high potential for creating alternative output pathways for the

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0308
http://psychologyinrussia.com

human brain. When support from passive BCIs becomes mature, the hybrid technology of the eye-brain-computer (EBCI) interface will have a chance to enable natural, fluent, and effortless interaction with machines in various fields of application.

Keywords: attention, eye-to-eye contact, eye movements, brain-computer interface (BCI), eye-brain-computer interface (EBCI), electroencephalography (EEG), expectancy wave (E-wave), human-robot interaction, brain output pathways

Introduction

No matter how rich our inner world, the intention to interact with objects in the external world or to communicate with others has to be implemented through the activity of peripheral nerves and muscles. Thus, motor system disorders, such as Amyothrophic Lateral Sclerosis (ALS) or a number of other diseases and traumas, seriously impair interaction with the external world, up to its full disruption in so-called Completely Locked-In Syndrome (CLIS). Of course, these clinical conditions are extremely rare. However, the problem of tight biological constraints on the motor output system frequently arises in the working life of many professionals, such as surgeons during an endoscopic operation or military servicemen on a battlefield: when both hands are busy with instruments, one suddenly wishes to possess something like a "third hand". Because of constraints imposed by the motor system's information transfer limits, the dependence of interaction on the motor system can be considered as a significant limitation in systems that combine natural human intellect and machine intelligence. This is akin to the philosophical stance of Henry Bergson (1907/2006), who emphasized the opposition of mental and physical efforts. Up to the present day, the effects of physical workload on mental activity have been little studied (DiDomenico, & Nussbaum, 2011). Nevertheless, we suspect that the physical load required for interaction with technologies can interfere with mental work, at least with kinds of mental work that require the highest concentration and creativity.

In this review, we consider basic elements of a technology that largely bypass the motor system and can be used for interaction with technical devices and people, either instead of normal mechanical or voice-based communication or in addition to it. More than two decades ago, one of us anticipated such aspects of the current development as the role of communication pragmatics for human-robot interaction (Velichkovsky, 1994) and the combination of eye tracking with brain imaging (Velichkovsky & Hansen, 1996). This direction is exactly where our work at Kurchatov Institute has been moving in recent years. In this review, we will focus primarily on our own studies, but discuss them in a more general context of contributions to the field by many researchers around the world. In particular, we will address requirements that are crucial for interaction with dynamic agents and not just with static devices. Even with simple geometric figures, motion dramatically enhances the observer's tendency to attribute intentionality (e.g., Heider & Simmel, 1944). This makes communication paradigms especially important in development of human-robot interaction systems.

Basic issues and approaches

Brain-Computer Interfaces (BCIs)

Many attempts have been made to bypass the usual pathway from the brain to a machine, based on manual activation of a keyboard, a mouse, a touchscreen, or any other type of mechanically controlled device. The most noticeable area of technology that has emerged as the result of such attempts is the Brain-Computer Interface (BCI), systems that offer fundamentally new output pathways for the brain (Brunner et al., 2015; Lebedev, & Nicolelis, 2017; Wolpaw, Birbaumer, McFarland, Pfurtscheller, & Vaughan, 2002). Among BCI studies, the most impressive achievements were based on the use of intracortically recorded brain signals, such as the simultaneous control by a patient for up to 10 degrees of freedom in a prosthetic arm (Collinger et al., 2013; Wodlinger et al., 2014). However, invasive BCIs are associated with high risks. The technology is yet not ready to be accepted even by severely paralyzed patients (Bowsher et al., 2016; Lahr et al., 2015; Waldert, 2016). This will possibly not change for years or even decades to come.

Noninvasive BCIs are based mainly on the brain's electric potentials recorded from the scalp, i.e., the electroencephalography (EEG). The magnetoencephalography (MEG) might offer a certain improvement over the EEG, but it is currently not practical due to its bulk and the fact that the equipment is very expensive. Similar difficulties arise with functional Magnetic Resonance Imaging (fMRI), which is too slow for communication and control of machines by healthy people; Near Infra-Red Spectroscopy (NIRS) is also too slow. Low speed and accuracy are currently associated with all types of noninvasive BCIs, and information transfer rate of even the fastest noninvasive BCIs is, unfortunately, much below the levels typical for the use of mechanical devices for computer control. Even a record typing speed for a non-invasive BCI, more than twice as fast as the previous record, still was only one character per second (Chen, Wang, Nakanishi, Gao, Jung, & Gao, 2015).

Non-invasive BCIs and eye movements

Even more important than the low speed and accuracy of most effective non-invasive BCIs is the fact that any substantial progress in performance has so far been associated only with BCIs that use visual stimuli. In such BCIs, the user is typically presented with stimuli repeatedly flashed on a screen at different positions associated with different commands or characters. In the BCI based on the P300 wave of event-related potentials (ERPs), flashes are presented at irregular intervals at time moments that differ for all positions or for different groups of positions. For example, if the BCI is used for typing, a matrix with flashing letters can be used. To type a letter, a user has to focus on it and notice its flashes; ERP to stimuli presented at different positions are compared, and the strongest response indicates the attended position. Many modifications to the P300 BCI have been proposed (Kaplan, Shishkin, Ganin, Basyul, & Zhigalov, 2013), whereby mostly aperiodic visual stimuli at different positions were used. Many P300 BCI studies involve healthy participants who automatically orient not only their focus of attention (a prerequisite for P300) but also their gaze toward the target. If such a participant is instructed to voluntarily refrain from looking at the stimulus (Treder, & Blankertz, 2010), or if a P300 BCI is used by a patient with impaired gaze control (Sellers, Vaughan, & Wolpaw, 2010), or if nonvisual modalities are used (Rutkowski, & Mori, 2015; Rutkowski, 2016), performance drops dramatically compared to the use of the P300 BCI when it is possible to foveate visual stimuli.

Another BCI technology is called the Steady-State Visual Evoked Potential BCI (SSVEP BCI; Gao, Wang, Gao, & Hong, 2014; Chen et al., 2015). It has relatively high accuracy and speed and employs periodic stimuli with different frequencies, which can in turn be found in the EEG response that these stimuli evoke. Other effective BCIs include the visual-ERP-based BCI developed in the mid-1970s (Vidal, 1973), the pseudorandom code-modulated Visual Evoked Potential BCI that uses complex temporal patterns of visual stimulation (cVEP BCI; Bin et al., 2011; Aminaka, Makino, & Rutkowski, 2015), and our "single-stimulus" BCI for rapidly sending a single command from only one position where a visual stimulus is presented aperiodically (Shishkin et al., 2013; Fedorova et al., 2014). In all these BCIs, responses of the visual cortex to visual stimuli are used explicitly. Although attention modulates these responses, and can be used separately from gaze (as in SSVEP BCI; see Lesenfants et al., 2014), the performance of these BCIs again drops greatly if foveating is prevented.

Enhancement of control by using eye movements contradicts the strict definition of a BCI as "a communication system in which messages or commands that an individual sends to the external world do not pass through the brain's normal output pathways of peripheral nerves and muscles" (Wolpaw et al., 2002). However, practical reason does not always stay within the limits of formal definitions. The oculomotor system is closely connected to the brain's attentional networks and is anatomically distinct from the main motor system (Parr, & Friston, 2017). In clinical conditions such as ALS, many paralyzed patients can control their gaze sufficiently well to benefit from gaze-based enhancement of BCIs. Moreover, gaze itself can be used, overtly and directly, to control computers, robots, and other machines, without any external stimulation and typically faster than with a BCI.

Gaze-based interaction

The main component of gaze-based interaction with technical systems and communication with other people is videooculography — eye tracking using a video camera that makes it possible to trace the position of the pupil and, based on that basis, to estimate the gaze coordinates. This technology is noninvasive and does not require attaching any sensors to the user's skin. Head movement restriction, once crucially important for obtaining high-quality data, is becoming less strict with the progress of technology, so a completely remote and non-constraining registration of eye movements is no longer unusual. Although under development for decades, in parallel to BCIs, eye-tracking technology was too expensive for use in consumer products, with the rare exceptions of communication systems for paralyzed persons (Pannasch, Helmert, Malischke, Storch, & Velichkovsky, 2008). Affordable eye trackers with sufficient capacities have recently appeared on the market, and applications of eye tracking are being developed for virtual and augmented reality helmets and even for smartphones, with the control and/or communication function considered as the most important. Jacob and Karn noted (2003, p. 589), "Before the user operates any mechanical pointing device, … the eye movement is available as an indication of the user's goal". This was demonstrated quantitatively in a number of studies; for example, it was shown that users tend to fixate on a display button or a link prior to approaching them manually or with the mouse cursor (Huang, White, & Buscher, 2012). It is not uncommon in interaction with computers that the mouse leads the gaze, but this is observed mainly for well-known locations (Liebling, & Dumais, 2014). Generally, visual fixations at an action location prior to the action are observed when objects in the physical world are being explored and manipulated (Johansson, Westling, Bäckström, & Flanagan, 2001; Land, Mennie, & Rusted, 1999; Neggers, & Bekkering, 2000; Velichkovsky, Pomplun, & Rieser, 1996).

In some practical applications, such as typing, gaze-based approaches became effective years ago (Bolt, 1982; Jacob, 1991; Velichkovsky, Sprenger, & Unema, 1997). Noninvasive BCI systems still cannot offer a level of speed, accuracy, and convenience similar to those of gaze typing systems. When targets are not too small, they can be selected using an eye tracker even faster than with a computer mouse (Ware, & Mikaelian, 1987; Sibert, & Jacob, 2000). A system can be tuned to respond to very short gaze dwells, e.g., 150–250 ms, so that the user gets a feeling of "a highly responsive system, almost as though the system is executing the user's intentions before he expresses them" (Jacob, 1991, p. 164). There is, however, a danger in such an extreme tuning, as it can lead to a vanishing sense of agency and abrupt deterioration of performance (Velichkovsky, 1995). Obviously an optimal threshold value has to be found in every particular case (Helmert, Pannasch, & Velichkovsky, 2008).

The Midas touch problem and natural gaze interaction

The most fundamental problem associated with eye-tracker-based interaction is known as the Midas touch problem: if an interface interprets visual fixation as a command, "you cannot look anywhere without issuing a command" (Jacob, 1991, p. 156). In the case of typing, areas outside a virtual keyboard can be made nonresponsive to gaze, so a user can simply avoid looking at the keyboard when he or she is not going to type; but it is difficult to avoid looking at the responsive area all the time. In a dynamic environment, e.g., in the case of robot control, or when the locations of response keys and areas for presenting visual information are close to each other, the Midas touch problem may make the interface annoying and inefficient. This is because the main function of gaze is to explore the visible environment, and this function normally is not under conscious control (Findlay, & Gilchrist, 2003). The eyes also move in an uncontrolled manner when we are thinking (Ehrlichman, & Micic, 2012; Walcher, Körner, & Benedek, 2017).

If interface is too responsive, its behavior becomes highly unnatural, because people "expect to be able to look at an item without having the look 'mean' something" (Jacob, 1991, p. 156). Almost all known means of avoiding the Midas touch problem require that the user move his or her eyes not naturally, but according to learned patterns (e.g., learned sequences of saccade directions or long dwells, e.g., 500 ms or longer), so the use of the interface often becomes tiresome and/or relatively slow (Majaranta, & Bulling, 2014). Jacob searched for patterns in noninstructed gaze behavior that can be used by the system to infer the user's goals. This led him to find that object selection using as short as 150-250 ms gaze dwell time works fine when selection can be easily undone (Jacob, 1991). Given that the eyes automatically got fixated for such a short time, very frequently without the intention to select anything, it was not surprising that it was he who coined the term "the Midas touch problem".

Is it possible to make eye movement input to the interface natural, and also the interface's response? Jacob (1993) noted that this is the case when interaction is organized similarly to how people respond to another person's gaze, although this approach is usually difficult to implement; for example, in one study, interaction was constructed in an analogy to a tour guide who estimates the visitor's interests by his or her gazes (Starker, & Bolt, 1990). It is likely that the basic function of gaze control which makes possible the use of gaze for interaction with machines is related not to vision, but to communication (Velichkovsky, Pomplun, & Rieser, 1996; Zhu, Gedeon, & Taylor, 2010), so it might be useful to learn more from communicative gaze behavior. This approach has attracted little attention over the decades of gaze interaction technology development, although it could lead to radical solutions of the Midas touch problem.

New developments at the Kurchatov Institute

Communicative gaze control of robots

Gaze alone — without speech, hand gestures, and (rarely, under natural conditions) without head movements — is used by humans to convey to other humans certain types of deictic information, mainly about spatial locations of interest. In everyday life, the role of this ability is, of course, far less prominent than the role of speech, but it can be comparable or more efficient for spatial information (Velichkovsky, 1995). The face and, especially, the eyes are powerful attractors of attention, stronger than the physical contrasts and semantic relations between the perceived objects (Velichkovsky et al., 2012). In particular, eye-to-eye contact is known to mobilize evolutionarily newer brain structures, an effect that can be observed even in humans who are facing an agent that is clearly virtual (Schrammel, Graupner, Mojzisch, & Velichkovsky, 2009). It was demonstrated that humans are very sensitive to a robot's gaze behavior, while they perfectly well realize that a robot is merely a machine: participants who were asked to judge a rescue robot's behavior felt more support from it when it "looked" at them (Dole et al., 2013) and emotional expression transfer from android avatars to human subjects was observed only in the case of simulated eye-to-eye contacts (Mojzisch, Schilbach, Helmert, Velichkovsky, & Vogeley, 2007).

An approach proposed for robot control by our group (Fedorova, Shishkin, Nuzhdin, & Velichkovsky, 2015; Shishkin, Fedorova, Nuzhdin, & Velichkovsky, 2014) was based on the developmental studies of joint attention, i.e., "simultaneous engagement of two or more individuals in mental focus on one and the same external thing" (Baldwin, 1995). The importance of joint attention was first noted by L.S. Vygotsky, and it remains an important focus of modern research (Carpenter, & Liebal, 2011; Tomasello, 1999). Importantly, joint attention gaze patterns are fast, and can function effectively even under high cognitive load (Xu, Zhang, & Geng, 2011). Last but not least, the Midas touch problem has not been observed with gaze communication in the joint attention mode, i.e., unintended eye movement caused by distractors or by lapses of attention do not normally lead to misinterpretation of information conveyed through gaze in this mode (Velichkovsky, Pomplun, & Rieser, 1996).

Although the human's ability to use eye movements for communication and control through eye-tracking technology has been addressed in the literature (Velichkovsky, 1995; Zhu, Gedeon, & Taylor, 2010), this understanding did not lead to making use of communicative gaze patterns in such a way that a machine is considered as a partner rather than a tool. We implemented joint attention patterns for control of a simple robot arm (R12-six, *ST Robotics*, UK). Because in two-way gaze communication a partner should have something that is perceived as eyes, with relevant "gaze behavior", a plain paper mask with "eyes" was attached to the robot's hand to provide it with certain anthropomorphic features (Fig. 1). The participant's eyes were tracked with a desktop eye tracker (Eyelink 1000 Plus, *SR Research*, Canada). The robot arm could be controlled by gaze patterns (here, predefined sequences of gaze fixations).

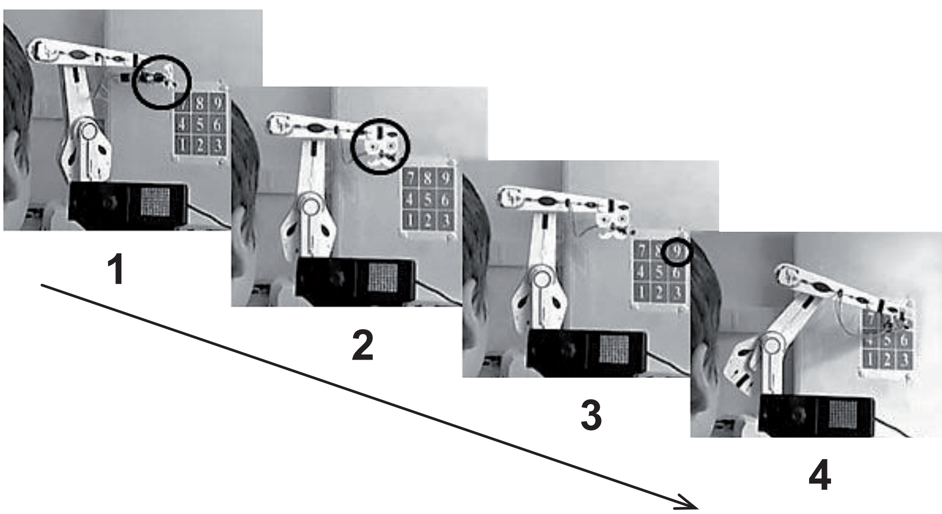


Figure 1. Gaze-based "communication" with a robot arm: a series of views from behind a participant (after Fedorova et al., 2015)

In the study, "communicative" patterns were compared with "instrumental" patterns. Specifically, "communicative" patterns were based on joint attention gaze patterns and included looking at the robot's "eyes" (see Figure 1). The pattern started from an "activating" gaze dwell of 500 ms or longer at the robot's "head" (1), and continued with the robot's turn toward the participant with resulting "eye-to-eye" contact (2). The immediately following human visual fixation location in the working field (3) was registered by the robot as the goal of the action required, so the robot pointed at the target with its "nose" as an emulated work instrument (4). In the "instrumental" pattern compared with this "communicative" pattern, a dwell on the "button" led to its lighting up and the robot's turning to a preparatory position, but without "looking" at the participant.

Participants were only told that the robot can be controlled using eye gaze and that they had to find a way, using their gaze only, to make the robot point at the target. Although not aware of what specifically they had to do with their eye movements, participants easily found both "communicative" and "instrumental" patterns, but showed no preference for either of them. However, it appeared from their reports that the robot's response to the "communicative" pattern was surprising and evoked the vivid impression that the robot shared their intention. This is distinct from what can be expected from a mechanical device; therefore, it may take time to get used to it. The study protocol included no practice and no testing of the hypothesis that "communicative" control can suppress gaze control's vulnerability to distractors and help to avoid the Midas touch problem. To decide whether it really offers significant benefits over the known strategies, further experiments are needed. Nevertheless, the present study confirmed the feasibility of "communicative" gaze control of robot behavior.

Selection of a moving target from a swarm

Interaction between a human operator and robot swarms has become an important area of human-machine interaction studies (Kolling, Nunnally, & Lewis, 2012). When an operator interacts with a large group of moving objects — in our case, the robots shown on a screen — selection of one of them with a mechanical pointer to receive detailed information from it or to send it a distinct command might not be an easy task, especially if they move in different directions, on different trajectories, and with varying speed. Fortunately, this situation is a special case of multiple objects tracking (MOT), fairly well investigated in cognitive science by Zenon Pylyshyn and his colleagues (e.g., Keane & Pylyshyn, 2006), who demonstrated that this task can be solved very quickly and in preattentive mode.

Following a moving object of interest may constitute a special case where interest in an object may be sufficient to select it by gaze in a natural way, without making artificially long static visual fixations. In this case, the gaze especially easily orients toward a moving object of interest and follows it continuously with high precision and without apparent effort, by using a distinct category of eye movements, the so-called dynamic visual fixations, or smooth pursuit (Brielmann, & Spering, 2015; Yarbus, 1967).

The use of smooth pursuit was recently considered to be of importance for interaction with technical devices (Esteves, Velloso, Bulling, & Gellersen, 2015). Its application to the robot selection problem was elaborated by our group (Zhao, Melnichuk, Isachenko, & Shishkin, 2017). In a preliminary study, robots were simulated on a computer screen by 20 balls; each of them was 2.5 deg in diameter and moving at a speed of 9 deg/s. Each ball had its own trajectory, changing direction each time it collided with other balls or with the screen's edges. By default, balls were gray, a target ball was indicated by red color, and selection changed the color to green. Selection was made with a cursor that was controlled, in different experiment conditions, either by a computer mouse or by gaze (an eye mouse). To select a ball, the cursor had to be closer to it than a threshold for 500 ms. Three different conditions of vibration-based feedback for a successful selection were used: without vibration, vibration in one or in two channels. With the computer mouse, the task required an average of about 1.7 s, while gaze-based selection was significantly faster, 1.1 s (Figure 2). Vibration feedback seemed to play no role in the selection efficiency.

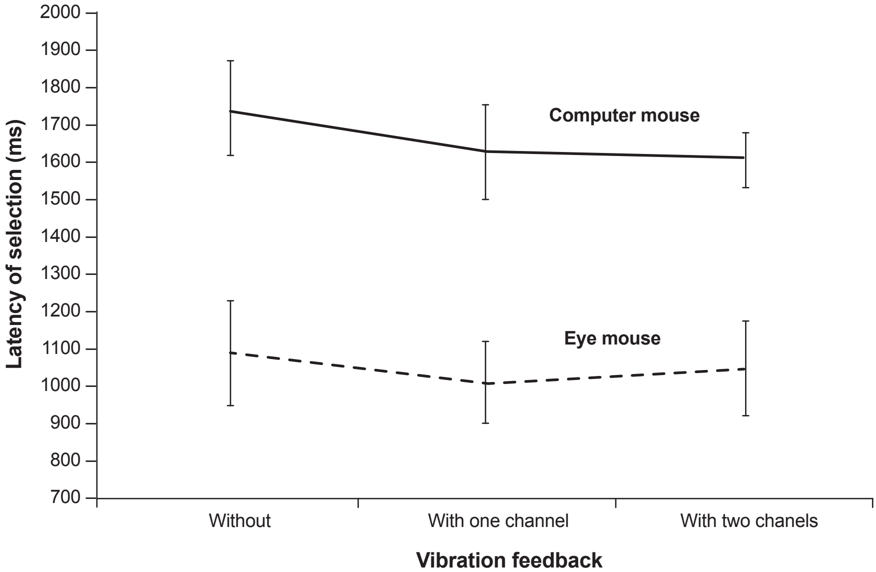


Figure 2. Time to selection of one of a number of moving targets in dependence on the output device and the vibration feedback

Perhaps perception of robots as animated autonomous agents is not necessary to enable interaction, as interest in the object and its movements may be sufficient to initiate and maintain smooth pursuit. In further studies, we will try to enhance moving robot selection using the EEG marker of intention, which is described in the next section. Another technical issue is that the selection of real objects in space needs a version of 3D binocular eye tracking, so algorithms for solving this classic measurement task have to be adopted (Wang, Pelfrey, Duchowski, & House, 2014; Weber, Schubert, Vogt, Velichkovsky, & Pannasch, 2017). In our current studies, we also aim at obtaining a shorter selection time by using advanced selection algorithms that were proposed for smooth-pursuit-based gaze interaction (Esteves et al., 2015).

Hunting for intention in the human brain

Velichkovsky and Hansen were the first authors who proposed solving the Midas touch problem by combining gaze-based control with a BCI: "point with your eye and click with your mind!" (Velichkovsky, & Hansen, 1996, p. 498). A number of research groups then tried to implement this idea by combining eye-tracker-based input with one of the existing BCIs, but the BCI component always added to the resulting hybrid system the worst features of non-invasive BCIs, namely low speed and accuracy, and the combination was not successful. A game-changing approach was proposed by Zander, based on his idea of "passive" BCIs that do not require the user's attention. Passive BCIs (Zander, & Kothe, 2011) monitor the user's brain state and react to it rather than to the user's explicit commands. In this approach, the BCI classified gaze fixations as spontaneous or intentional (Ihme, & Zander, 2011; Protzak, Ihme, & Zander, 2013), presumably by using an EEG response related to expectation of the gaze-controlled interface feedback, or, paradoxically, expectation is used to trigger the action that is expected.

The most relevant EEG phenomenon was discovered about 50 years ago by Grey Walter. He used an experimental paradigm in which a warning stimulus preceded an imperative stimulus (one requiring a response) with a fixed time interval between them. This component, a slow negative wave, was called the Contingent Negative Variation (CNV). As early as 1966, Walter proposed, in an abstract for the EEG Society meeting, that the expectancy wave (E-wave), the non-motor part of CNV, "can be made to initiate or arrest an imperative stimulus directly, thus by-passing the operant effector system" (Walter, 1966, p. 616). The E-wave, or the Stimulus-Preceding Negativity (SPN) — this latter name was introduced for a slightly different experimental paradigm (Brunia, & Van Boxtel, 2001) — is what can be expected to appear in the gaze fixations intentionally used for interaction.

Zander's group did not study this EEG marker in detail, and their experimental design included only visual search. In this condition, the P300 wave appears in fixations on targets, allowing the researcher to differentiate target and non-target fixations by this, very different EEG component (Kamienkowski, Ison, Quiroga, & Sigman, 2012; Brouwer, Reuderink, Vincent, van Gerven, & van Erp, 2013; Ušćumlić, & Blankertz, 2016). While the combination of the P300 and gaze fixation also may be a promising tool for human-machine interaction, it cannot be used for sending commands to machines deliberately. In addition, dwell time used in these studies was too long (1 s, in Protzak, Ihme, & Zander, 2013). Therefore, we designed a study where gaze dwells with a shorter threshold (500 ms) were used to trigger freely chosen actions.

In our research (Shishkin et al., 2016; Velichkovsky et al., 2016), spontaneous and intention-related gaze dwells were collected when the participants played *EyeLines*, a gaze- controlled version of the computer game *Lines*. In *EyeLines*, each move consists of three gaze dwells: (1) switching the control on by means of a gaze dwell at a remote "switch-on" location, (2) selection of one of the balls presented in the game field, (3) dwell on a free cell to which the ball had to be moved. After the ball was moved and before the next fixation on the "switch-on" location, no fixation had any effect, so that spontaneous fixations could be collected. Special efforts were made to ensure that eye-movement-related electrophysiological artifacts did not affect the analyzed EEG intervals. In all participants, a negative wave was indeed discovered in the gaze dwells used for control and it was absent or had lower amplitude in the spontaneous fixations. The results are shown in Figure 3. Note that the waveforms reveal no signs of P300. Based on statistical features extracted from 300 ms EEG intervals (200-500 ms relative to dwell start), intentional and spontaneous dwells could be classified on a single-trial basis with accuracy much greater than the random level (Shishkin et al., 2016).

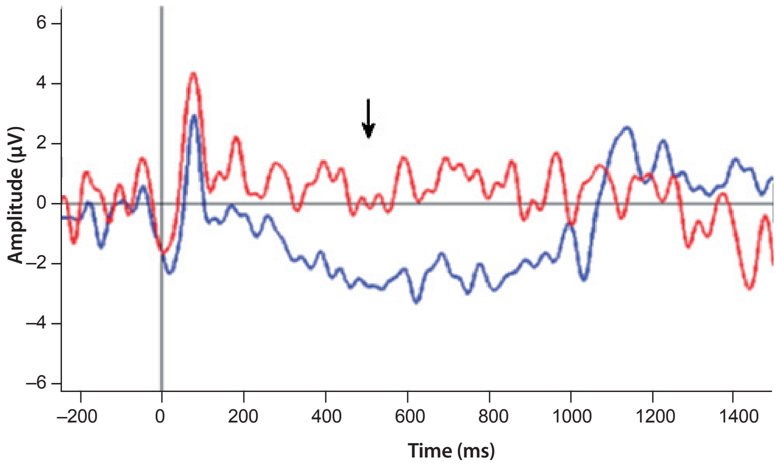


Figure 3. Fixation-related brain potentials (POz, grand average, n = 8) for gaze dwells intentionally used to trigger actions are shown by the blue line and for spontaneous fixations by the red line. "0" corresponds to dwell start; visual feedback was presented at 500 ms, corresponding to the arrow position (after Shishkin et al., 2016; Velichkovsky et al., 2016)

In subsequent studies, we found that the EEG marker for the gaze dwells intentionally used for control does not depend on gaze direction (Korsun et al., 2017) and demonstrated that initially chosen approaches to construct feature sets and the classifier can be further improved (Shishkin et al., 2016a). Preliminary attempts to classify the 500 ms gaze dwells online using the passive expectation-based BCI (Nuzhdin et al., 2017a; Nuzhdin et al., 2017 in press) so far have not shown a significant improvement compared to gaze alone. This could be related to a suboptimal classifier and/or inadequate choice of tests, because in the game we used, actions were quickly automated at the beginning of our research, while SPN amplitude is likely to decrease precisely under such conditions. We are now improving the classifier and preparing better tests to investigate in detail the capacity and limitations of the eye-brain-computer interface (EBCI), as we call this new hybrid system. If this line of applied cognitive research is successful, it could result in interfaces responding to the user's intentions more easily and without annoying, unintended activations.

Conclusion

We discussed here ways to bypass the common limits of the brain motor system in human-machine interaction. Invasive BCIs can be surprisingly efficient, but their use may remain too risky for decades to come. Noninvasive BCIs are also not well suited for this purpose, and their progress in performance by healthy users is critically dependent on gaze's impact upon the selection of spatial locations. In our overview of the new developments at the Kurchatov Institute, we discussed the communicative nature of human-robot interaction and approaches to build a more efficient technology on this basis. Specifically, "communicative" patterns of interaction can be based on joint attention paradigms from developmental psychology, and including a mutual exchange eye-to-eye "looks" between human and robot. Further, we provided an example of eye mouse superiority over the computer mouse, here in emulating the task of selecting a moving robot from a swarm. Finally, a passive noninvasive BCI that uses EEG correlates of intention was demonstrated. This may become an important filter to separate intentional gaze dwells from non-intentional ones. These new approaches show a high potential for creating alternative output pathways for the human brain. When support from passive BCIs matures, the hybrid ECBI technology will have a chance to enable natural, fluent, and effortless interaction with machines in various fields of application.

According to Douglas Engelbart (1962), intellectual progress often depends on reducing the efforts needed for interaction with artificial systems. Centered on this idea, he developed basic elements of the modern human-computer interface, such as the mouse, hypertext, and the elements of the graphical user interface. These tools have radically improved humans' interaction with computers and partially with robots, but they have not completely excluded physical efforts from the interaction process. Could we become even more effective in solving intellectual tasks, at least when mental concentration is crucial, if our collaboration with technical devices were totally free of physical activity? Further studies are needed to answer this question, but cognitive interaction technologies seem to be becoming advanced enough to conduct such experiments in the near future.

Acknowledgments

This study was in part supported by the Russian Foundation for Basic Research (RFBR ofi-m grants 15-29-01344). We wish to thank E.V. Melnichuk for his participation in the experiments.

References

Allison, B., Luth, T., Valbuena, D., Teymourian, A., Volosyak, I., & Graser, A. (2010). BCI demographics: How many (and what kinds of) people can use an SSVEP BCI? *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 18*(2), 107–116. doi: 10.1109/TNSRE.2009.2039495

Aminaka, D., Makino, S., & Rutkowski, T.M. (2015). Chromatic and high-frequency cVEP-based BCI paradigm. *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE* (pp. 1906–1909). IEEE Press. doi: 10.1109/EMBC.2015.7318755

Baldwin, D.A. (1995). Understanding the link between joint attention and language. In C. Moore & P. J. Dunham (Eds.). *Joint attention: Its origins and role in development* (pp. 131–158). Hillsdale, NJ: Lawrence Erlbaum.

Bergson, H. (2006). *L'Évolution créatrice*. Paris: PUF. (Original work published 1907)

Bin, G., Gao, X., Wang, Y., Li, Y., Hong, B., & Gao, S. (2011). A high-speed BCI based on code modulation VEP. *Journal of Neural Engineering*, *8*(2), 025015. doi: 10.1088/1741-2560/8/2/025015

Bolt, R.A. (1982). Eyes at the interface. *Proceedings of 1982 Conference on Human Factors in Computing Systems* (pp. 360–362). New York: ACM Press. doi: 10.1145/800049.801811

Bolt, R.A. (1990). A gaze-responsive self-disclosing display. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3–10). New York: ACM Press.

Bowsher, K., Civillico, E.F., Coburn, J., Collinger, J., Contreras-Vidal, J.L., Denison, T., ... & Hoffmann, M. (2016). Brain–computer interface devices for patients with paralysis and amputation: A meeting report. *Journal of Neural Engineering*, *13*(2), 023001. doi: 10.1088/1741-2560/13/2/023001

Brielmann, A.A., & Spering, M. (2015). Effects of reward on the accuracy and dynamics of smooth pursuit eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(4), 917–928. doi: 10.1037/a0039205

Brouwer, A.M., Reuderink, B., Vincent, J., van Gerven, M.A., & van Erp, J.B. (2013). Distinguishing between target and nontarget fixations in a visual search task using fixation-related potentials. *Journal of Vision*. 13:17. doi: 10.1167/13.3.17

Brunia, C.H. M., & Van Boxtel, G.J.M. (2001). Wait and see. *International Journal of Psychophysiology, 43*, 59–75. doi: 10.1016/S0167-8760(01)00179-9

Brunner, C., Birbaumer, N., Blankertz, B., Guger, C., Kübler, A., Mattia, D., ... & Ramsey, N. (2015). BNCI Horizon 2020: Towards a roadmap for the BCI community. *Brain-Computer Interfaces*, *2*(1), 1–10.

Carpenter, M., & Liebal, K. (2011). Joint attention, communication, and knowing together in infancy. In A. Seemann (Ed.), *Joint attention: New developments in psychology, philosophy of mind, and social neuroscience* (pp. 159–181). Cambridge, MA: MIT Press.

Chen, X., Wang, Y., Nakanishi, M., Gao, X., Jung, T.P., & Gao, S. (2015). High-speed spelling with a noninvasive brain–computer interface. *Proceedings of the National Academy of Sciences*, *112*(44), E6058–E6067. doi: 10.1073/pnas.1508080112

Collinger, J.L., Wodlinger, B., Downey, J.E., Wang, W., Tyler-Kabara, E.C., Weber, D. J., … & Schwartz, A.B. (2013). High-performance neuroprosthetic control by an individual with tetraplegia. *The Lancet*, *381*(9866), 557–564. doi: 10.1016/S0140-6736(12)61816-9

DiDomenico, A., & Nussbaum, M.A. (2011). Effects of different physical workload parameters on mental workload and performance. *International Journal of Industrial Ergonomics*, *41*(3), 255–260. doi: 10.1016/j.ergon.2011.01.008

Dole, L.D., Sirkin, D.M., Currano, R.M., Murphy, R.R. & Nass, C.I. (2013). Where to look and who to be: Designing attention and identity for search-and-rescue robots. *Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction* (pp. 119–120). IEEE Press. doi: 10.1109/HRI.2013.6483530

Engelbart, D.C. (1962). *Augmenting human intellect: A conceptual framework*. SRI Summary Report AFOSR-322.

Ehrlichman, H., & Micic, D. (2012). Why do people move their eyes when they think? *Current Directions in Psychological Science*, *21*(2), 96–100. doi: 10.1177/0963721412436810

Esteves, A., Velloso, E., Bulling, A., & Gellersen, H. (2015). Orbits: Gaze interaction for smart watches using smooth pursuit eye movements. *Proc. 28th Annual ACM Symposium on User Interface Software & Technology* (pp. 457–466). New York: ACM Press. doi: 10.1145/2807442.2807499

Fedorova, A.A., Shishkin, S.L., Nuzhdin, Y.O., Faskhiev, M.N., Vasilyevskaya, A.M., Ossadtchi, A.E., . . . Velichkovsky, B.M. (2014). A fast "single-stimulus" brain switch. In G. Muller-Putz et al. (Eds.), *Proceedings of 6th International Brain-Computer Interface Conference* (Article 052). Graz: Verlag der Technischen Universität Graz. doi:10.3217/978-3-85125-378-8-52

Fedorova, A.A., Shishkin, S.L., Nuzhdin, Y.O., & Velichkovsky, B.M. (2015). Gaze based robot control: The communicative approach. In *7th International IEEE/EMBS Conference on Neural Engineering (NER)*. (pp. 751–754). IEEE Press.

Findlay, J.M., & Gilchrist, I.D. (2003). *Active vision: The psychology of looking and seeing.* Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780198524793.001.0001

Gao, S., Wang, Y., Gao, X., & Hong, B. (2014). Visual and auditory brain–computer interfaces. *IEEE Transactions on Biomedical Engineering*, *61*(5), 1436–1447. doi: 10.1109/TBME.2014.2300164

Guger, C., Allison, B.Z., Großwindhager, B., Prückl, R., Hintermüller, C., Kapeller, C., . . . & Edlinger, G. (2012). How many people could use an SSVEP BCI? *Frontiers in Neuroscience*, *6*, 169. doi: 10.3389/fnins.2012.00169

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, *57*(2), 243–259. doi: 10.2307/1416950

Helmert, J.R., Pannasch, S. & Velichkovsky, B.M. (2008). Influences of dwell time and cursor control on the performance in gaze driven typing. *Journal of Eye Movement Research*, *2*(4), 3, 1–8.

Huang, J., White, R., & Buscher, G. (2012). User see, user point: Gaze and cursor alignment in web search. *Proceedings of SIGCHI Conference on Human Factors in Computing Systems* (pp. 1341–1350). New York: ACM Press. doi: 10.1145/2207676.2208591

Ihme, K., & Zander, T.O. (2011). What you expect is what you get? Potential use of contingent negative variation for passive BCI systems in gaze-based HCI. *Affective computing and intelligent interaction* (ACII 2011), 447–456. doi: 10.1007/978-3-642-24571-8_57

Jacob, R.J. (1991). The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems (TOIS)*, *9*(2), 152–169. doi: 10.1145/123078.128728

Jacob, R.J. (1993). Eye movement-based human-computer interaction techniques: Toward non-command interfaces. *Advances in Human-Computer Interaction*, *4*, 151–190.

Jacob, R.J., & Karn, K.S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In R. Radach, J. Hyona, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573–605). Oxford, UK: Elsevier. doi: 10.1016/B978-044451020-4/50031-1

Johansson, R.S., Westling, G., Bäckström, A., & Flanagan, J.R. (2001). Eye–hand coordination in object manipulation. *Journal of Neuroscience*, *21*(17), 6917–6932.

Kamienkowski, J.E., Ison, M.J., Quiroga, R.Q., & Sigman, M. (2012). Fixation-related potentials in visual search: A combined EEG and eye tracking study. *Journal of Vision, 12*, 4. doi: 10.1167/12.7.4

Kaplan, A.Y., Shishkin, S.L., Ganin, I.P., Basyul, I.A., & Zhigalov, A.Y. (2013). Adapting the P300-based brain–computer interface for gaming: A review. *IEEE Transactions on Computational Intelligence and AI in Games*, *5*(2), 141–149. doi: 10.1109/TCIAIG.2012.2237517

Keane, B., & Pylyshyn, Z.W. (2006). Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function. *Cognitive Psychology*, *52*(4), 346–368. doi: 10.1016/j.cogpsych.2005.12.001

Kolling, A., Nunnally, S., & Lewis, M. (2012). Towards human control of robot swarms. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction* (pp. 9–96). New York: ACM Press. doi: 10.1145/2157689.2157704

Korsun, O. V., Shishkin, S.L., Medyntsev, A.A., Fedorova, A.A., Nuzhdin, Y.O., Dubynin, I.A., & Velichkovsky, B.M. (2017). The EEG marker for the intentional gaze dwells used to interact with a computer is independent of horizontal gaze direction. *Proceedings of the 4th Conference "Cognitive Science in Moscow: New Studies"*, 428–433.

Lahr, J., Schwartz, C., Heimbach, B., Aertsen, A., Rickert, J., & Ball, T. (2015). Invasive brain–machine interfaces: A survey of paralyzed patients' attitudes, knowledge and methods of information retrieval. *Journal of Neural Engineering*, *12*(4), 043001. doi: 10.1088/1741-2560/12/4/043001

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, *28*(11), 1311–1328. doi: 10.1068/p2935

Lebedev, M.A., & Nicolelis, M.A. (2017). Brain-Machine Interfaces: From basic science to neuroprostheses and neurorehabilitation. *Physiological Reviews*, *97*(2), 767–837. doi: 10.1152/physrev.00027.2016

Lesenfants, D., Habbal, D., Lugo, Z., Lebeau, M., Horki, P., Amico, E., … & Laureys, S. (2014). An independent SSVEP-based brain–computer interface in locked-in syndrome. *Journal of Neural Engineering*, *11*(3), 035002. doi: 10.1088/1741-2560/11/3/035002

Liebling, D.J., & Dumais, S.T. (2014). Gaze and mouse coordination in everyday work. *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication* (pp. 1141–1150). New York: ACM Press. doi: 10.1145/2638728.2641692

Majaranta, P., & Bulling, A. (2014). Eye tracking and eye-based human–computer interaction. In S.H. Fairclough and K. Gilleade (Eds.), *Advances in physiological computing: Human–computer interaction series*. (pp. 39–65). London: Springer-Verlag. doi: 10.1007/978-1-4471-6392-3_3

Manyakov, N.V., Chumerin, N., & Van Hulle, M.M. (2012). Multichannel decoding for phase-coded SSVEP brain–computer interface. *International Journal of Neural Systems*, *22*(05), 1250022. doi: 10.1142/S0129065712500220

Mojzisch, A., Schilbach, L., Helmert, J.R., Velichkovsky, B.M., & Vogeley, K. (2007). The effects of self-involvement on attention, arousal, and facial expression during social interaction with virtual others. *Social Neuroscience, 1*(1), 184–195.

Neggers, S.F., & Bekkering, H. (2000). Ocular gaze is anchored to the target of an ongoing pointing movement. *Journal of Neurophysiology*, *83*(2), 639–651.

Nuzhdin, Y.O., Shishkin, S.L., Fedorova, A.A., Trofimov, A.G., Svirin, E.P., Kozyrskiy, B.L., … Velichkovsky, B.M. (2017a). The expectation-based Eye-Brain-Computer Interface: An attempt of online test. *Proceedings of the 2017 ACM Workshop on An Application-Oriented Approach to BCI Out of the Laboratory* (pp. 39–42). New York: ACM Press. doi: 10.1145/3038439.3038446

Nuzhdin, Y.O., Shishkin, S. L., Fedorova, A.A., Kozyrskiy, B.L., Medyntsev, A.A., Svirin, … Velichkovsky, B.M. (2017 in press). Passive detection of feedback expectation: Towards fluent hybrid eye-brain-computer interfaces. *Proceedings of the 7th International Brain-Computer Interface Conference.* Graz: Verlag der Technischen Universität Graz.

Pannasch, S., Helmert, J.R., Malischke, S., Storch, A., & Velichkovsky, B.M. (2008). Eye typing in application: A comparison of two systems with ALS patients. *Journal of Eye Movement Research, 2*(4):6, 1–8.

Parr, T., & Friston, K.J. (2017). The active construction of the visual world. *Neuropsychologia*, *104*, 92–101. doi: 10.1016/j.neuropsychologia.2017.08.003

Protzak J., Ihme, K., & Zander, T.O. (2013). A passive brain-computer interface for supporting gaze-based human-machine interaction. *International Conference on Universal Access in Human-Computer Interaction* (pp. 662–671). Heidelberg: Springer. doi: 10.1007/978-3-642-39188-0_71

Rutkowski, T.M. (2016). Robotic and virtual reality BCIs using spatial tactile and auditory oddball paradigms. *Frontiers in Neurorobotics*, *10*, 20. doi: 10.3389/fnbot.2016.00020

Rutkowski, T.M., & Mori, H. (2015). Tactile and bone-conduction auditory brain computer interface for vision and hearing impaired users. *Journal of Neuroscience Methods*, *244*, 45–51. doi: 10.1016/j.jneumeth.2014.04.010

Schrammel, F., Graupner, S.-T., Mojzisch, A., & Velichkovsky, B.M. (2009). Virtual friend or threat? The effects of facial expression and gaze interaction on physiological responses and emotional experience. *Psychophysiology, 46*(5), 922–931. doi: 10.1111/j.1469-8986.2009.00831.x

Sellers, E.W., Vaughan, T.M., & Wolpaw, J.R. (2010). A brain-computer interface for long-term independent home use. *Amyotrophic Lateral Sclerosis*, *11*(5), 449–455. doi: 10.3109/17482961003777470

Shishkin, S.L., Fedorova, A.A., Nuzhdin, Y.O., Velichkovsky, B.M. (2014). Upravleniye robotom s pomoschyu vzglyada: kommunikativnaya paradigma [Gaze-based control in robotics: The communicative paradigm.] In B.M. Velichkovsky, V.V. Rubtsov, D.V. Ushakov (Eds.), *Cognitive Studies, 6,* 105–128.

Shishkin, S.L., Ganin, I. P., Basyul, I.A., Zhigalov, A.Y., & Kaplan, A.Y. (2009). N1 wave in the P300 BCI is not sensitive to the physical characteristics of stimuli. *Journal of Integrative Neuroscience*, *8*(4), 471–485. doi: 10.1142/S0219635209002320

Shishkin, S.L., Fedorova, A.A., Nuzhdin, Y.O., Ganin, I.P., Ossadtchi, A.E., Velichkovsky, B.B., … Velichkovsky, B.M. (2013). Na puti k vysokoskorostnym interfeysam glaz-mozg-kompyuter: sochetaniye "odnostimulnoy" paradigm I perevoda vzglyada [Toward high-speed eye-brain-computer interfaces: Combining the single-stimulus paradigm and gaze reallocation]. *Vestnik Moskovskogo Universiteta. Seriya 14: Psikhologiya [Moscow University Psychology Bulletin]*, *4*, 4–19.

Shishkin, S.L., Kozyrskiy, B. L., Trofimov, A.G., Nuzhdin, Y.O., Fedorova, A.A., Svirin, E.P., & Velichkovsky, B.M. (2016a). Improving eye-brain-computer interface performance by using electroencephalogram frequency components. *RSMU Bulletin, 2*, 36–41.

Shishkin, S.L., Nuzhdin, Y.O., Svirin, E.P., Trofimov, A.G., Fedorova, A.A., Kozyrskiy, B.L., & Velichkovsky, B.M. (2016b). EEG negativity in fixations used for gaze-based control: Toward converting intentions into actions with an Eye-Brain-Computer Interface. *Frontiers in Neuroscience*, *10*, 528. doi: 10.3389/fnins.2016.00528

Sibert, L.E., & Jacob, R.J.K. (2000). Evaluation of eye gaze interaction. In *Proceedings of the ACM CHI 2000 Human Factors in Computing Systems Conference,* 281–288. New York: Addison-Wesley/ACM Press. doi: 10.1145/332040.332445

Tomasello, M. (1999). *The cultural origins of human cognition*. New York: Harvard University Press.

Treder, M.S., & Blankertz, B. (2010). (C) overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and Brain Functions*, *6*(1), 28. doi: 10.1186/1744-9081-6-28

Ušćumlić, M., & Blankertz, B. (2016). Active visual search in non-stationary scenes: coping with temporal variability and uncertainty. *Journal of Neural Engineering,* 13:016015. doi: 10.1088/1741-2560/13/1/016015

Velichkovsky, B.M. (1994). Towards pragmatics of human-robot interaction (Keynote address to the Congress). In J. Arnold, & J.M. Prieto (Eds.), *Proceeding of the 23rd International Congress of Applied Psychology*, 12–15. Washington, DC: APA Press.

Velichkovsky, B.M. (1995). Communicating attention: Gaze position transfer in cooperative problem solving. *Pragmatics and Cognition*, *3*(2), 199–224. doi: 10.1075/pc.3.2.02vel

Velichkovsky, B.M., Nuzhdin, Y.O., Svirin, E. P., Stroganova, N.A., Fedorova, A.A., & Shishkin, S.L. (2016). Upravlenie silij mysli: Na puti k novym formam vzaimodejstvija cheloveka s technicheskimi ustrojstvami [Control by "power of thought": On the way to new forms of interaction between human and technical devices]. *Voprosy Psikhologii [Issue of Psychology]*, *62*(1), 79–88.

Velichkovsky, B.M., Cornelissen, F., Geusebroek, J.-M., Graupner, S.-Th., Hari, R., Marsman, J.B., & Shevchik, S.A. (2012). Measurement-related issues in investigation of active vision. In B. Berglund, G.B. Rossi, J. Townsend, & L. Pendrill (Eds.), *Measurement with persons: Theory and methods* (pp. 281–300). London/New York: Taylor and Francis.

Velichkovsky, B.M., & Hansen, J.P. (1996). New technological windows into mind: There is more in eyes and brains for human-computer interaction (Keynote address to the Conference). *Proceedings of ACM CHI-96: Human factors in computing systems* (pp. 496–503). New York: ACM Press.

Velichkovsky, B.M., Pomplun, M., & Rieser. H. (1996). Attention and communication: Eye-movement-based research paradigms. In W.H. Zangemeister, S. Stiel & C. Freksa (Eds.), *Visual attention and cognition*, 125–254. Amsterdam/New York: Elsevier. doi: 10.1016/S0166-4115(96)80074-4

Velichkovsky, B.M., Sprenger, A., & Unema, P. (1997). Towards gaze-mediated interaction: Collecting solutions of the "Midas touch problem". In S. Howard, J. Hammond, and G. Lindgaard (Eds.), *Human-Computer Interaction INTERACT'97* (pp. 509–516). London: Chapman and Hall. doi: 10.1007/978-0-387-35175-9_77

Vidal, J.J. (1973). Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering*, *2*(1), 157–180. doi: 10.1146/annurev.bb.02.060173.001105

Waldert, S. (2016). Invasive vs. non-invasive neuronal signals for brain-machine interfaces: Will one prevail? *Frontiers of Neuroscience, 10*, 295. doi: 10.3389/fnins.2016.00295

Walcher, S., Körner, C., & Benedek, M. (2017). Looking for ideas: Eye behavior during goal-directed internally focused cognition. *Consciousness and Cognition*, *53*, 165–175. doi: 10.1016/j.concog.2017.06.009

Walter, W.G. (1966). Expectancy waves and intention waves in the human brain and their application to the direct cerebral control of machines. *Electroencephalography and Clinical Neurophysiology, 21*, 616.

Wang, R.I., Pelfrey, B., Duchowski, A.T., & House, D.H. (2014). Online 3D gaze localization on stereoscopic displays. *ACM Transactions on Applied Perception, 11*(1), 1–21. doi: 10.1145/2593689

Ware, C., & Mikaelian, H.T. (1987). An evaluation of an eye tracker as a device for computer input. *Proceedings of the ACM CHI+GI'87 Human Factors in Computing Systems Conference,* 183–188. New York: ACM Press.

Weber, S., Schubert, R.S., Vogt, S., Velichkovsky, B.M., & Pannasch, S. (2017 in press). Gaze3DFix: Detecting 3D fixations with an ellipsoidal bounding volume. *Behavior Research Methods*.

Wodlinger, B., Downey, J.E., Tyler-Kabara, E.C., Schwartz, A.B., Boninger, M.L., & Collinger, J.L. (2014). Ten-dimensional anthropomorphic arm control in a human brain–machine interface: Difficulties, solutions, and limitations. *Journal of Neural Engineering*, *12*(1), 016011. doi: 10.1088/1741-2560/12/1/016011

Wolpaw, J.R., Birbaumer, N., McFarland, D.J., Pfurtscheller, G., & Vaughan, T.M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, *113*(6), 767–791. doi: 10.1016/S1388-2457(02)00057-3

Xu, S., Zhang, S., & Geng, H. (2011). Gaze-induced joint attention persists under high perceptual load and does not depend on awareness. *Vision Research*, *51*(18), 2048–2056. doi: 10.1016/j.visres.2011.07.023

Yarbus, A.L. (1967). *Eye movements and vision*. New York: Plenum Press. doi: 10.1007/978-1-4899-5379-7

Zander, T.O., & Kothe, C. (2011). Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. *Journal of Neural Engineering*, *8*(2), 025005. doi: 10.1088/1741-2560/8/2/025005

Zhao, D.G., Melnichuk, E. V., Isachenko, A.V., & Shishkin, S.L. (2017). Vybor robota iz mobilnoy gruppirovki s pomoschyu vzglyada [Gaze-based selection of a robot from a mobile group]. *Proceedings of the 4th Conference "Cognitive Science in Moscow: New Studies"*, 577–581.

Zhu, D., Gedeon, T., & Taylor, K. (2010). Head or gaze? Controlling remote camera for hands-busy tasks in teleoperation: A comparison. *Proceedings of the 22nd Conference Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction*, 300–303. New York: ACM Press. doi: 10.1145/1952222.1952286

Original manuscript received May 16, 2017
Revised manuscript accepted June 11, 2017
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society · Lomonosov Moscow State University

Structure of conceptual models in the senior operating staff of nuclear power plants

Aleksandr A. Oboznov[a*], Elena D. Chernetskaya[b], Yulia V. Bessonova[a]

[a] *Institute of Psychology, the Russian Academy of Sciences, Moscow, Russia*
[b] *"ROSATOM — CICE&T"*

[*] Corresponding author. E-mail: aao46@mail.ru

Background. The relationships between conceptual model structures and an operator's professional efficiency are of direct practical importance, particularly in the case of large-scale industrial complexes combining several human-machine systems. A typical example is the power unit of a nuclear power plant (NPP).

Objective and Methods. The purpose of this study was to explore the conceptual models of senior reactor operators (SROs) of NPPs. The study involved 64 men working as SRO at five NPPs in Russia. The methods included: structured interviews, expert estimations, multidimensional scaling (ALSCAL), the K-means clustering algorithm, and frequency analysis. The procedure was as follows: 32 key characteristics of the power unit were defined, including shift operators' jobs and duties, technical subsystems, types of equipment, and the crucial power unit parameters. The participants were offered a 32×32 matrix for pair-wise estimation of the strength of the links between these key characteristics on a seven-point scale (496 links in total).

Results. A general scheme of key characteristics in the conceptual models was defined. This scheme was displayed in the operators regardless of their employment history. Within the scheme, however, two types of conceptual models were identified, which could be distinguished by the relative number of strong links between the key characteristics. With respect to intersystem links including key characteristics of the reactor and turbine NPP departments, this number was significantly higher in models of Type 1 than in those of Type 2. A positive correlation between the number of these links and the professional efficiency indicators was also established. Operators with Type 1 models were able to more predictably represent the power unit operation.

Conclusion. The main role in creating predictable and efficient conceptual models was played by strong intersystem links in mental representations of workflow.

Keywords: Nuclear Power Plant (NPP) senior reactor operator (SRO), conceptual model, mental image, multidimensional scaling, workflow, subjective strength of links, professional efficiency

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0309
http://psychologyinrussia.com

Introduction

L.S. Vygotsky (1982) stated that the development of mind causes changes not so much in existing mental functions, as in the links between them, resulting in the emergence of new psychological systems. This observation is fully applicable to the development of new cognitive structures necessary to regulate professional activity. One of these is the "conceptual model" used by human operators in their work. The notion was introduced into human-factor engineering and psychology by A.T. Welford (1961) as a workflow "mental image" necessary for the operator to control the process. This model, although inexact, allows the operator to correlate different workflow stages, and to act proactively and efficiently. In the Russian literature, this notion has become quite widespread (Zinchenko, 1970; Lomov, 1977; Galaktionov, 1992; Oboznov, 2009; Oboznov et al., 2013a; Bessonova, 2012; Ryabov, 2014; Chernetskaya, 2014). In fact, it is similar to one of the basic concepts of Russian psychology, the "operative image" (Oshanin, 1969).

A study of conceptual models proceeds from a scientific description of their functions to the identification of their content and structures. The *content* consists of the operator's views and knowledge of the human — machine system (HMS) workflows, working environment, typical problem situations, decision-making rules, operator's tasks and methods, programs of control (Munipov, & Zinchenko, 2001), as well as the required workflow dynamics (Oshanin, 1977). The content of a conceptual model also includes knowledge gained in professional experience. Of course, the representations of knowledge in a conceptual model has to be organized in a certain way, a *structure*. The reason that cognitive science is interested in mental structures is to better understand how the acquisition, memorization, transfer, and use of new information depends on the ways that pre-existing knowledge is organized (Anderson, 2002; Velichkovsky, 2006).

This also applies to the HMS operators' conceptual model structures. The results of a few empirical studies have shown that, for skilled operators, the structures of conceptual models differ little from one another; at the same time, they differ greatly from the ways that novices in the profession organize workflow-related knowledge (Cooke, & McDonald, 1987; Golikov, 2003). The relationships between conceptual model structures and an operator's professional efficiency are of direct practical importance, particularly in large-scale industrial complexes combining several HMSs. A typical example is the power unit of a nuclear power plant (NPP). The power unit includes two HMSs, a reactor department and a turbine department, which, in turn, include a number of subsystems. To understand the power unit operation, predict trends, and make decisions, especially in abnormal situations, the conceptual model of senior reactor operators (SROs) should represent both the intrasystem links among the characteristics of the reactor department and the intersystem links among the characteristics of the reactor and turbine departments. The difference between these types of links is that the intrasystem links, due to their limited mutual influences, are more predictable, while the intersystem links are much less predictable due to the complexity of their direct and indirect interaction (Golikov, & Costin, 1999). This creates an additional hazard potential of such HMS complexes for society and the environment. The present study was carried out on a large group of SROs to identify their types of conceptual models and, in particular, the relationships between the structure of those conceptual models and the operators' professional efficiency.

Method

Participants

The study involved 64 men working as SROs at five Russian NPPs. Their length of service ranged from 6 months to 34 years.

Procedure

The study included four stages. At the first stage, two experts (NPP operators with high professional qualification and over 10 years of work experience) specified a list of key characteristics of a power unit to be contained in the SRO conceptual models. At the second stage, the participants gave a subjective assessment of the strength of links between the key characteristics of the power unit. At the third stage, a general scheme of the SROs' conceptual model structures and their types was defined. At the fourth and final stage, the SROs' professional efficiency was evaluated and its relationship with the type of conceptual model was determined.

Structured interviews were used to identify the key characteristics of the power unit to be contained in the SRO conceptual models. To estimate the strength of links between the key characteristics, participants were provided with a 32×32 matrix for pair-wise comparison of the indicated characteristics. The strength of links between all the key characteristics (496 in total) was estimated on a seven-point scale: 7 = very strong; 6 = strong; 5 = above average; 4 = average; 3 = below average; 2 = weak; 1 = very weak. Further, these links were analyzed in three aspects. In the first, the links were considered by the criterion of their subjective strength for the SROs, namely, as strong, average, or weak. In the second aspect, these links were considered by the criterion of their affiliation, that is, as intra-system ones, linking the key characteristics only to the "Reactor Department" HMS or the "Turbine Department" HMS, and intersystem ones, linking the key characteristics of both HMSs. In the third aspect, the same links were considered by the criterion of their function in the SRO's professional activity. The following links were analyzed:

- links that perform a *cognitive* function and reflect the operation of the power unit as a technical complex without the operator's direct intervention (power unit operation links), that is, among the technical subsystems, equipment, aggregates and their parameters; a set of SRO representations of these links is the *cognitive* component of the conceptual model;
- links that perform a *regulatory* function, i.e., between the operators' duties and the controlled technical subsystems, units of equipment, aggregates and their parameters (power unit control links); these links reflect the zones of each operator's personal responsibility for the management and control of a certain power unit section; a set of SRO representations of these links is the *regulatory* component of the conceptual model;
- links that perform a *communicative* function, i.e., between the duty shift operators' jobs and duties (operator communication links); these links reflect the power unit operators' interactions; a set of SRO representations of these links is the *communicative* component of the conceptual model.

To define a general structural scheme of the conceptual models, a 32 × 32 matrix was used, averaged over the entire group of 64 participants. This matrix underwent multidimensional scaling (ALSCAL), followed by the construction of a two-dimensional semantic space. The conceptual model structural types were identified using the K-means clustering algorithm.

Finally and in order to determine the SRO's professional efficiency, the expert estimation method was used. Expert estimations were made on a 9-point scale by a group of three experts: the immediate supervisor of the operator being evaluated, a training center instructor, and a full-time psychologist engaged in emergency training of operators on a full-scale power unit simulator. The consistency of expert estimations was determined using Kendall's coefficient of concordance (W).

Results

The structured interview with two experts revealed four classes of key characteristics of power units, which should be contained in the SRO conceptual models:

- shift operators' jobs (NPP shift supervisor, power unit shift supervisor, operator-inspectors, etc.);
- shift operators' duties (operational personnel management, operational control of the primary circuit parameters, etc.);
- power unit technical subsystems and equipment (vacuum system, main circulation pump, turbine generator, etc.);
- parameters of power unit technical subsystems and equipment (power reactivity margin, electrical capacity, etc.).

In total, for further research, the experts selected 32 key characteristics (8 characteristics of each type).

The data presented in Table 1 shows how the SRO's understanding of the strength of links depended on the function of these links in professional activity.

According to the reports, strong and very strong (6–7 points) power unit operation links (cognitive function) meant that, when one key characteristic changes, another key characteristic will always (or almost always) change too. In fact, a strong link was understood by the operators as a cause-and-effect relationship. The average (3–5 points) power unit operation links meant that, if the condition of one technical subsystem (parameter) changes, the condition of another technical subsystem (parameter) might or might not change with equal probability. The weak (1–2 points) power unit operation links meant that, that, if the condition of one technical subsystem (parameter) changes, the condition of another technical subsystem (parameter) changes very rarely, if ever.

The strong and very strong (6–7 points) power unit control links (regulatory function) meant that the SRO official duties are related to their personal responsibility for the management and control of particular technical subsystems, aggregates and their parameters. The average (3–5 points) power unit control links meant that the SRO's official duties are related to their indirect responsibility for the management and control of particular technical subsystems, aggregates and their

Table 1. SRO understanding of the strength of links between key characteristics of the power unit

Functions	SRO understanding of the strength of links between the key characteristics of the power unit
Cognitive (power unit operation links)	*Strong and very strong (6–7 points)*: when one characteristic changes, it will (almost) always cause a change in another characteristic (cause-and-effect relationship). *Average (3–5 points)*: when one characteristic changes, another one might or might not change, with approximately equal probability. *Weak and very weak (1–2 points)*: when one characteristic changes, another characteristic changes very rarely, if ever.
Regulatory (power unit control links)	*Strong and very strong (6–7 points)*: the operator bears personal (direct) responsibility for the management and control of a particular type of equipment or its parameter. *Average (3–5 points)*: the operator bears indirect responsibility for the management and control of a particular type of equipment or its parameter. *Weak and very weak (1–2 points)*: the operator bears no responsibility for the management and control of a particular type of equipment or its parameter.
Communicative (operator communication links)	*Strong and very strong (6–7 points)*: joint involvement in activities occurs (almost) always. *Average (3–5 points)*: joint involvement in activities occurs or does not occur with equal probability. *Weak and very weak (1–2 points)*: joint involvement in activities occurs very rarely, if ever.

parameters, which were not areas of the SRO's personal responsibility, but could be indirectly influenced by their actions. The weak (1–2 points) power unit control links meant that operators' official duties were unrelated to the management and control of particular technical subsystems and their parameters.

The strong and very strong (6–7 points) operator communication links (communicative function) meant that one operator could perform his official duties only if another operator did so. The average (3–5 points) operator communication links meant that the dependence of one operator's performing his official duties on another operator's doing so might or might not be manifested, according to the circumstances. The weak (1–2 points) operator communication links meant that one operator's performing his official duties did not depend on another operator's doing so.

The common point in the SROs' understanding of the strength of all the considered links is as follows: the stronger these links, the more predictable the mutual influences of the key characteristics being linked. At the same time, the relative number of strong links was significantly different in the cognitive, regulatory, and communicative components of the conceptual model (see Table 2).

The data presented in Table 2 suggest that the power unit operation links (cognitive component) were estimated by the senior reactor operators mostly as weak. Generally speaking, this representation adequately reflects the NPP power unit operation as a human–machine complex which has numerous intra- and intersystem links, including non-linear and unstable interactions, unstable and extreme work-

Table 2. Distribution (%) of links of different strengths in the conceptual model components (data averaged over the total group of 64 senior reactor operators)

Components	Strength of links			Total (%)
	Weak	Average	Strong	
Cognitive component (power unit operation links)	48.0	24.0	28.0	100.0
Regulatory component (power unit control links)	32.0	26.5	41.5	100.0
Communicative component (operator communication links)	22.5	19.5	58.0	100.0

ing conditions, the influence of subjective factors associated with the operational personnel control input, etc. (Anokhin, & Ostreykovsky, 2001). A different picture was observed in the power unit control links (regulatory component). These links were estimated by the operators as strong in 41.5% of cases, or 1.5 times more often than in the power unit operation links. The operator communication links (communicative component) were estimated as strong even more often, in 58% of cases. This means that for the SROs, the most predictable were the links reflecting the duty shift operators' interactions.

As a result of applying multidimensional scaling to the 32 × 32 matrix averaged for all 64 participants, a generalized semantic space of the conceptual model was

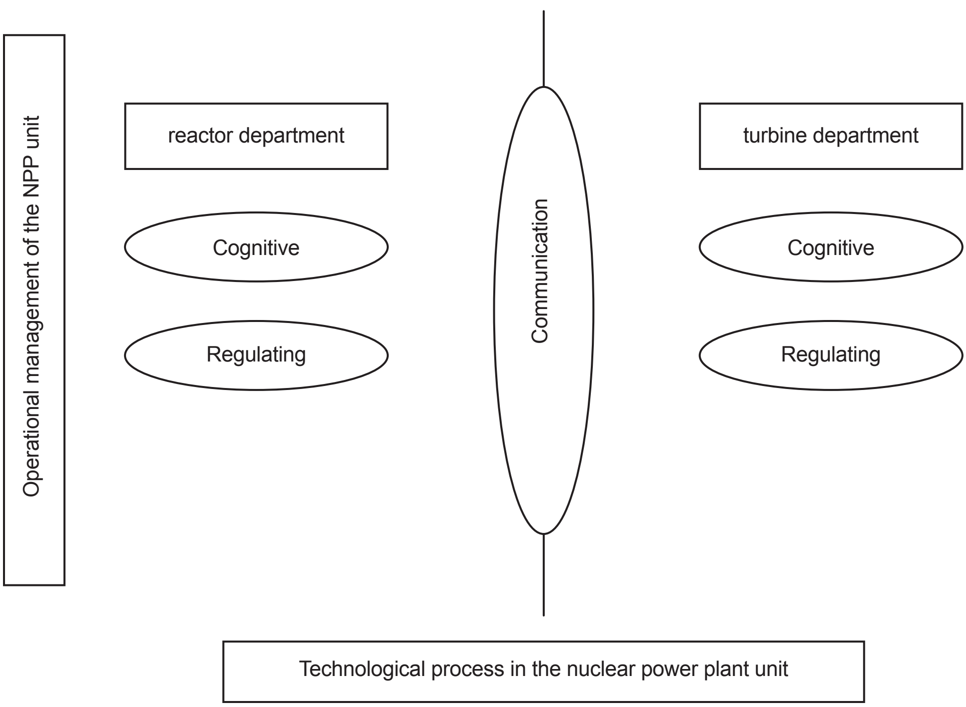


Figure 1. General structural scheme of SRO conceptual model

obtained, which included 32 key characteristics of the power unit. Based on the analysis of the resulting semantic space, a general structural scheme of the SRO conceptual model was drawn up (see Figure 1).

Along the horizontal axis are the power unit operation links (cognitive component) and unit control links (regulatory component). The reactor department operation and control links are followed by those of the turbine department. This sequence reflects the main technological process control in the power unit: reactor-assisted coolant production (using water heated to a certain temperature), and coolant transformation into steam energy and then into electric power in the turbine generator. For this reason, the horizontal axis of the semantic space is interpreted as the "Technological process in the nuclear power plant unit" factor. Along the vertical axis, there are the links between the duty shift operators (communicative component), reflecting the SRO's views of the duty shift operators' interactions. Placed in the lower part of the vertical axis are the links between the characteristics reflecting the executive operators' jobs and duties; in the upper part are the links between the characteristics reflecting the supervisory duty shift operators' jobs and duties. The vertical axis is interpreted as the "Operational management of the NPP unit" factor. This general scheme suggests that the main technological process in the power unit and its control by the duty shift operators are presented in the SROs' conceptual model as relatively independent processes.

On the basis of the K-means clustering analysis, we came to a conclusion about the existence of two types of the power unit conceptual models, whereby both types retain the same general scheme (see Figure 2).The criteria for distinguishing these types were the relative number of strong and weak intra- and intersystem links, as well as the links between the operators. As Figure 3 shows, the relative number of strong (6–7 points) links of each kind in Type 1 conceptual models is significantly (1.9 ÷ 2.1) larger than in Type 2 ($p \leq 0{,}01$; F-criterion). On the contrary, the relative number of weak (1–2 points) links of each kind in Type 1 conceptual models is considerably smaller than in Type 2 ($p \leq 0{,}01$; F-criterion).

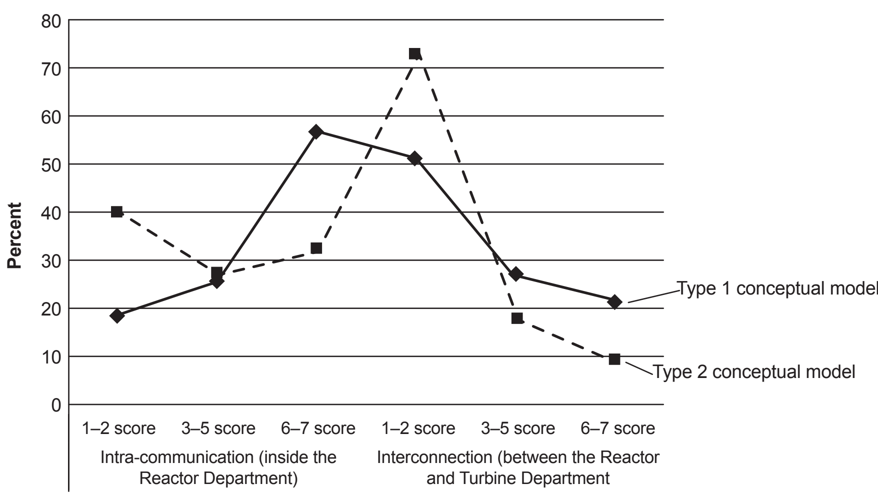


Figure 2. Two types of SRO conceptual models

The distinction of these two types of conceptual models by this criterion was confirmed by the summarized data on distribution of links of different subjective strength (see Table 3). The relative number of strong (6–7 points) links between the key characteristics of the power unit in Type 1 conceptual models was 3.2 times larger, 41.0%, as against 13.0% in Type 2 ($p < 0.01$; chi-squared test).

Table 3. Distribution (%) of links of different subjective strength depending on the conceptual model type

Conceptual model type	Subjective strength of links			Total (%)
	Weak	Average	Strong	
Type 1	12.0	47.0	41.0	100.0
Type 2	29.0	58.0	13.0	100.0

To illustrate this point, individual semantic spaces for each of the two conceptual model types are presented in Figures 3 and 4. These semantic spaces demonstrate the structural specifics of both conceptual model types. The subjective strength of links between the key characteristics is expressed in the distance between them: the stronger the link, the closer the key characteristics.

Figure 3 shows an individual semantic space of the Type 1 power unit conceptual model. Five subgroups include the key characteristics of the reactor and

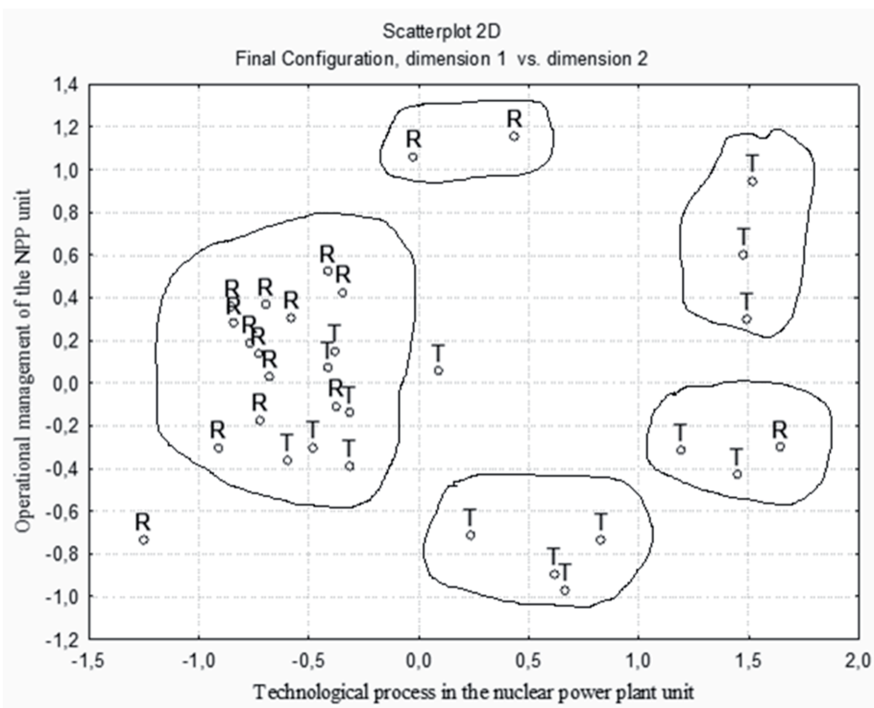


Figure 3. Individual semantic space of an SRO with Type 1 conceptual model (R, T = key characteristics of the reactor and turbine departments, respectively)

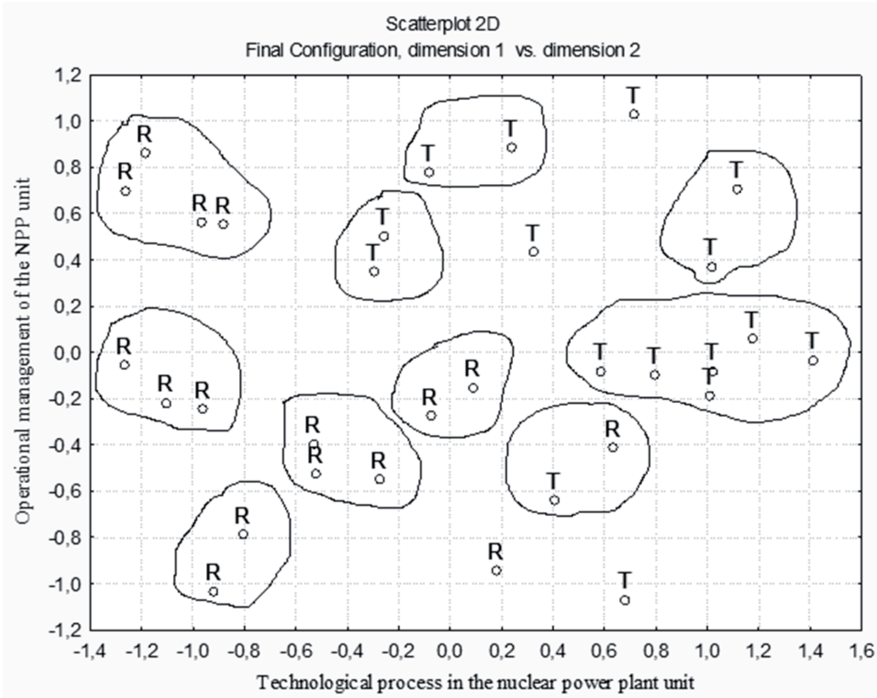


Figure 4. Individual semantic space of an SRO with Type 2 conceptual model (Nomenclatures are the same as in Fig. 3)

turbine departments (these subgroups were defined according to the clustering analysis, Ward's method, and city-block distance). The largest subgroup (Figure 3, left) contains 18 key characteristics, including 12 characteristics of the reactor department and 6 characteristics of the turbine department. Inclusion of these characteristics into one subgroup meant that the senior reactor operator considered the links between them strong. By their nature, these links were both intra- and intersystem ones. This conceptual model allowed the senior reactor operator to create a holistic and predictable representation of the operation of not only the reactor department for which he was responsible, but of the entire power unit, and to make independent decisions in regular and abnormal situations. No wonder that this SRO had the highest rate of professional efficiency of all the 64 participants**.**

Figure 4 shows an SRO individual semantic space with a Type 2 power unit conceptual model. In this model, 10 subgroups of key characteristics were identified, i.e., twice that of the previous case. As a rule, the subgroups contained 2–3 key characteristics only of the reactor or the turbine departments. Only two subgroups included the key characteristics of both departments. That is, in this conceptual model, strong intrasystem links dominated, while a small number of strong intersystem links prevented the creation of a holistic and predictable representation of the power unit operation. As a result, this SRO often was unable to make independent decisions and had the lowest rate of professional efficiency of all the 64 participants.

Discussion

Here we address the main question of this study, namely the relationship between the structure of a conceptual model and the operator's profession efficiency. According to the results of professional efficiency scaling, the SROs were distinguished as having high, average, or low efficiency. A qualitative criterion for distinguishing the professional efficiency levels was the operators' ability to perform job duties in regular as well as in hazardous situations. Thus, the SROs with high professional efficiency included those who were able to independently perform their duties under both regular and abnormal conditions; those with average professional efficiency could independently perform their duties in regular situations, but in emergencies a supervisor should monitor their work; and those with low professional efficiency made mistakes in regular situations and needed significant additional training to improve their qualification (this category of persons is not recommended for the operator profile positions).

The results of the study show a clear relationship between the conceptual model types and SRO professional efficiency. The estimations made by the group of three experts were highly consistent: the values of Kendall's coefficient of concordance (W) were in the range of 0.79–0.87. The evidence for a relationship between the conceptual model types and SRO professional efficiency was twofold. First, differences were revealed in the conceptual model types of senior reactor operators with high and average professional efficiency. The operators with high professional efficiency had predominantly Type 1 conceptual models with an average of 8 subgroups of key characteristics, while those with average professional efficiency had Type 2 conceptual models averaging 12 subgroups of key characteristics. It can be assumed that the operators with high professional efficiency, due to the larger number of highly interrelated key characteristics contained in the conceptual models, create a more holistic representation of the power unit operation.

The second type of evidence demonstrating the relationship between the conceptual model types and SRO professional efficiency was the positive correlation found between the level of professional competency and the number of strong *intersystem* links (Spearman's rho = 0.24; $p < 0.05$). This means that with an increase in the number of strong intersystem links, in transition from conceptual models of Type 2 to those of Type 1, the SRO's professional efficiency increased. This correlation between professional efficiency and the number of strong intersystem links, together with the absence of such a correlation for intrasystem links, point to a special role of strong intersystem links for the SROs to envisage the power unit operation as a whole, supporting their ability to make independent decisions.

Conclusion

The SRO conceptual models contain representations about the links between the key characteristics of the power unit that differ in their functions, connections, and subjective strength. This latter parameter indicates the predictable mutual influences of these characteristics: the stronger the links, the more predictable these interactions. For the operators, the most predictable were the links reflecting their interaction, and the least predictable were the power unit operation links between its technical subsystems and parameters. A general structural scheme of the key characteristics of the power unit can be explicated and presented as a two-dimensional semantic space, where the power unit operation and control links are along the horizontal axis, and the links between the duty shift operators are along the vertical axis. Two types of the power unit conceptual model were distinguished according to the relative number of strong intra- and intersystem links between the key characteristics of the power unit. In Type 1 conceptual models, this number is definitely larger than in Type 2 conceptual models. For the operators with Type 1 conceptual models, the power unit operation was more predictable than for those with Type 2 models. Strong intersystem links played the most important part in creating more predictable representations. A positive, although relatively low, correlation was found between the number of strong intersystem links in conceptual models and SRO professional efficiency indicators. We speculate that this accounts for their ability to make independent decisions in risky situations.

The revealed links between the SRO conceptual model types and professional efficiency and personality traits suggest the following areas where the obtained results may be applied:

– selection of operators and evaluation of their psychological readiness to independently manage HMS in regular situations and in emergencies;
– development of intelligent support systems for forming the operators' conceptual models required to manage HMS.

The results of an earlier study (Oboznov et al., 2013b) demonstrated the practical feasibility of this latter task.

Acknowledgements

The study was supported by the Russian Humanitarian Science Foundation, grants 08-06-00403a and 11-06-01136a.

References

Anderson, J. (2002).*Cognitivnaya psikhologiya* [Cognitive psychology]. Saint Petersburg, Russia: Piter.

Anokhin, A.N., & Ostreykovsky, V.A. (2001). Voprosy ergonomiki v yadernoy energetike [Ergonomic issues in the nuclear industry]. Moscow: Energoatomizdat.

Bessonova, Yu.V. (2012). Typologiya mentalnykh reprezentatsiy professionalnykh zadach u spetsialistov atomnoy energetiki [Typology of job task mental representations for nuclear reactor operators]. *Eksperimentalnaya psihologiya [Experimental Psychology]*, *5*(2), 102–118.

Chernetskaya, E.D. (2014). Tipy kontseptualnykh modeley u operatorov atomnykh stantsiy [Conceptual model types of nuclear power plant operators]. In A.N. Anokhin, P.I. Paderno, & S.F. Sergeyev (Eds.), *Trudy Mezhdunarodnoy nauchno-prakticheskoy konferentsii "Psikhologiya truda, inzhenernaya psichologiya i ergonomika, 2014 (Ergo-2014)"* [Proceedings of the international conference "Psychology of labor, engineering psychology and ergonomics, Ergo-2014"] (pp. 194-201). Saint Petersburg, July 3–5, 2014. Saint Petersburg, Russia: Mezhregionalnaya ergonomicheskay assotsiatsiya.

Cooke, N.M., & McDonald, J.E. (1987).The application of psychological scaling techniques to knowledge elicitation for knowledge-based systems. *International Journal of Man-Machine Studies*, *26*, 533–550. doi: 10.1016/S0020-7373(87)80086-X

Galaktionov, A.I. (1992). Sistemnoe issledovanie psychicheskikh obrazov, formiruemykh operatorom-tekhnologom [System research of operators' mental images]. In V.A. Bodrov & V.F. Venda (Eds.), *Sistemnyy podkhod v inzhenernoy psikhologii i psikhologii truda* [Systemic approach to engineering psychology and psychology of labor]. (pp. 92–105). Moscow: Nauka.

Golikov, Yu.Ya. (2003). *Metodologiya psihologicheskikh problem proektirovaniya tekhniki* [Methodology of psychology in engineering design]. Moscow: PER SE.

Golikov, Yu.Ya., & Kostin, A.N. (1999). Teoriya i metody analiza problemnostey v slozhnoy operatorskoy deyatelnosti [Theory and method for problem analysis of complex action]. In *Problemnost v professionalnoy deyatelnosti* [Issues in professional activity]. Moscow: Institute of Psychology RAS.

Lomov, B.F. (1977). O putyakh postroyeniya teorii inzhenernoy psikhologii na osnove sistemnogo podkhoda [Ways to construct a theory of engineering psychology on the basis of a systematic approach]. In *Inzhenernaya psikhologiya. Teoriya, metodologiya, prakticheskoye primeneniye* [Engineering psychology. Theory, methodology and application]. (pp. 31–55). Moscow: Nauka.

Munipov, V.M., & Zinchenko, V.P. (2001). *Ergonomika: cheloveko-orientirovannoe proektirovanie tekhniki, programmnykh sredstv i sredy* [Ergonomics: Human-oriented design for equipment, software and working environment]. Moscow: Logos.

Oboznov, A.A. (2009). Struktura konceptualnoy modeli u cheloveka-operatora: sistemnyy podkhod [The conceptual model structure of the human operator: A systemic approach]. In V.A. Bodrov & A.L. Guravlev (Eds.), *Aktualnye problemy psikhologii truda, inzhenernoy psikhologii i ergonomiki* [Topical issues in psychology of labor, engineering psychology and ergonomics]. (Vol. 1, pp. 403–413). Moscow: Institute of Psychology RAS.

Oboznov, A.A., Chernetskaya, E.D., & Bessonova, Yu.V. (2013a). Kontseptualnye modeli atomnoy stantsii u operatorov s raznym professionalnym stazhem [Conceptual models of nuclear power plant operators with various amounts of professional experience]. *Psychologicheskiy zhurnal [Psychological Journal], 4*(34), 47–57.

Oboznov, A.A., Dozortsev, V.M., Nazin, V.A., & Gutsykova, S.V. (2013b). Intellektualnaya systema formirovaniya u operatorov mentalnykh reprezentatsiy tekhnologicheskogo obyekta [An intelligent system for constructing mental images of technological infrastructure]. *Rossiyskiy nauchnyy zhurnal [Russian Scientific Journal]*, *7*(38), 133–138.

Oshanin, D.A. (1969). Rol operativnogo obraza v vyyavlenii informatsionnogo soderzhaniya signalov [The role of operative images in revealing the information content of signals]. *Voprosy psikhologii [Issues in Psychology]*, *4*, 24–34.

Oshanin, D.A. (1977) Kontseptsiya operativnosti otrazheniya v obshey i inzhenernoy psikhologii [The concept of the operativeness of reflection in general and engineering psychology]. In *Inzhenernaya psikhologiya: Teoriya, metodologiya, prakticheskoe primenenie* [Engineering psychology. Theory, methodology and application]. (pp. 134–149). Moscow: Nauka.

Ryabov, V.B. (2014). Kontseptualnye modeli i ikh ispolzovaniye v prikladnykh inzhenerno-psikhologicheskikh i ergonomicheskikh issledovaniyakh [Conceptual models and their use in applied psychological and ergonomic research]. In A.A. Oboznov, A.L. & Zhuravlev (Eds.), *Aktualnye problem psikhologii truda, inzhenernoy psihologii i ergonomiki* [Topical issues in psychology of labor, engineering psychology and ergonomics]. (Vol. 6, pp. 285–298). Moscow: Institute of Psychology RAS.

Velichkovsky, B.M. (2006). *Cognitivnaya nauka: Osnovy psikhologii poznaniya* [Cognitive science: Foundations of epistemic psychology]. Vol. 2. Moscow: Smysl/Academia.

Vygotsky, L.S. (1982) O psychologicheskikh sistemakh. In: A.R. Luria & M.G. Yarochevsky (Eds.) *Vygotsky L.S. Sobranie sochineniy v 6 tomakh.*Vol.1. Voprosy teorii i istorii psihologii [Collected works in 6 volumes. Vol. 1. Theory and history of psychology] (pp. 109–131). Moscow: Pedagogika.

Welford, A.T. (1961). On the human demands of automation: Mental work conceptual model, satisfaction and training. *Industrial and Business Psychology, 5*, 182–193.

Zinchenko, V.P. (1970). Analiz deyatelnosti operatora [Analysis of the operator's action]. In *Ergonomika: printsipy i rekomendatsii* [Ergonomics: Principles and recommendations]. Vol. 1. Moscow: VNIITE.

Original manuscript received June 01, 2017
Revised manuscript accepted July 22, 2017
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society
Lomonosov Moscow State University

The application of virtual reality technology to testing resistance to motion sickness

Galina Ya. Menshikova, Artem I. Kovalev,
Oxana A. Klimova, Valentina V. Barabanschikova

Faculty of Psychology, Lomonosov Moscow State University, Moscow, Russia

* Corresponding author. E-mail:

Background. Prolonged exposure to moving images in virtual reality systems can cause virtual reality induced motion sickness (VIMS). The ability to resist motion sickness may be associated with the level of vestibular function development. **Objective.** The aim of the present research is to study the oculomotor characteristics of individuals whose observation of moving virtual environments causes the VIMS effect. We hypothesized that people who have a robust vestibular function as a result of their professional activity, are less susceptible to VIMS than people who have no such professional abilities. The differences in people's abilities to resist the effects of the virtual environment may be revealed in the oculomotor characteristics registered during their interaction with a virtual environment.

Design. Figure skaters, football players, wushu fighters, and non-trained people were tested. The CAVE virtual reality system was used to initiate the VIMS effect. Three virtual scenes were constructed consisting of many bright balls moving as a whole around the observer. The scenes differed in the width of the visual field; all balls subtended either 45°, 90° or 180°.

Results. The results showed more active eye movements for athletes compared to non-trained people, i.e. an increase in blink, fixation, and saccade counts. A decrease in saccadic amplitudes was revealed for figure skaters. These characteristics were considered specific indicators of the athletes' ability to resist motion sickness.

Conclusions. It was found that the strength of the VIMS effect increased with the increasing width of the visual field. The effectiveness of virtual reality and eye-tracking technologies to test the VIMS effect was demonstrated.

***Keywords*:** virtual reality technology, motion sickness, vestibular dysfunction, vection illusion, eye movement characteristics, professional abilities of athletes

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0310
http://psychologyinrussia.com

Introduction

During the last decade, virtual reality technologies became a common method used in various psychological studies. The advantages and challenges of applying these systems in experimental psychology, organizational psychology, and sports psychology have been thoroughly described and analyzed (Zinchenko, Menshikova, Bayakovskiy, Chernorizov & Voiskounsky, 2010). The perspective for using these technologies has also been shown to play a profound role at diverse levels of innovative learning and teaching, including professional education.

Virtual reality technology is a system of visualization tools, which includes special virtual environment devices (CAVE-systems, systems of augmented reality, HMD helmets, spherical displays), as well as simpler devices such as widescreen projection 3D-displays, 3D-theaters, and virtual reality goggles. Even in the first stages of using these technologies for the development of different skills (mostly in training the spatial abilities of pilots and astronauts), negative symptoms were found, which included vertigo, nausea, spatial disorientation, and problems with balance and movement execution. Moreover, many observers noted that they had a strong impression of having moved their bodies during the interaction with virtual environment, although objectively their bodies remained stationary (Hettinger, 2002). This phenomenon has been called the self-motion illusion, or the vection illusion.

The above-mentioned negative symptoms were initially attributed to technical flaws in the virtual reality technologies themselves (Biocca, 1992). However, it has been shown that the improvement of technical characteristics (higher video resolution, and more accurate tracking and optical systems) leads to increased symptoms: for example, the latency of their appearance is significantly reduced (Bailey, Denis, Goldsmith, Hall & Sherwood, 1994).

The complex of negative symptoms which appear in virtual environments, was initially associated with motion sickness, which occurs in the natural environment in people who travel on ships or aircraft. It was suggested that motion sickness appears because of the conflict between the sensory signals from vestibular and visual systems (Reason, 1978). For example, a person standing in a ship cabin sees the stationary surroundings of the cabin (no sense of body movement through the visual signals), while he/she perceives body movement through vestibular signals. However, later on, to describe feelings of discomfort arising from interaction with virtual reality systems, new terms were proposed, which reflected new symptoms which can be compared with those caused by commonly used devices such as centrifuges. These terms frequently correspond to the type of virtual reality device: "simulator sickness" from aviation, car, or motorcycle simulators (Kellogg, Kennedy & Graybiel, 1964); "cyber-sickness" from videogames on big screens (Keshavarz & Berti, 2014); and "cinema-sickness" from 3D-theaters (Griffin, 2012).

Other studies showed a stronger influence of the virtual environment: of 1102 participants in one experiment, 142 (12.9%) stopped the experiment, and 960 (87.1%) reported a high level of discomfort (Sharples, Cobb, Moody & Wilson, 2008). The symptoms described were so robust, that in 2005 the International Organization for Standardization (ISO) compelled the manufacturers of widescreen displays to conduct special tests of the potential for inducing vection illusion in users before the product could be put on the market.

To solve these problems in the use of virtual reality devices, it is necessary to both assess their impact, and to identify the individual characteristics of the people who are able to resist their impact. In previous studies, both objective and subjective procedures were developed to find appropriate measurements. Large individual differences in discomfort symptoms have been revealed, so methods of subjective assessment became the most common. One subjective method is the generally accepted Simulator Sickness Questionnaire (SSQ). It was developed by R.S. Kennedy and his colleagues (Kennedy, Lane, Kevin, Berbaum & Lilienthal, 1993) on the basis of the Pensacola motion sickness questionnaire (MSQ). The latter was created by NASA to assess the condition of prospective astronauts after their training in centrifuge and swimming pool weightlessness simulators (Kellogg, Kennedy & Graybiel, 1964). Three results were found–nausea, oculomotor reactions, and disorientation–which were used to calculate a total score indicating the severity of the simulator sickness.

Using the MSQ and SSQ measures, studies showed that simulator sickness occurred in only some of the study subjects, and its intensity changed depending on the level of subjects' adaptation to the virtual environment. This data allowed researchers to suggest that different people can resist the impact of virtual reality devices by relying on special skills they have obtained in their professional activities (McLeod, Reed, Gilson & Glennerster, 2008). This suggestion has been confirmed in experiments (Howarth & Costello, 1997) which tested the negative impact of helmets constructed by different manufacturers. The SSQ results were estimated for participants who played a shooter computer game for an hour. It appeared that their simulator sickness severity depended more on skills of interaction with the virtual games than on helmet's technical specifications. Specifically, those participants who had more experience playing computer games, especially car and aviation simulators, scored lower points on the SSQ scale.

The influence of user mobility on the severity of the simulator sickness was also demonstrated in studies of user-control in a virtual environment (Stanney, Hale, Nahmens, & Kennedy, 2003). It was shown that in situations where participants were allowed to actively move around in a virtual environment, they had lower rates of discomfort: only a few subjects complained of simulator disorders. The successful participants were found to have had a great deal of past experience interacting with computer games, so it took less time for them to adapt to user control activities in the virtual environment.

However, the MSQ and SSQ measurements have some disadvantages: first, the estimates are based on memories of past virtual events; second, SSQ measurements do not allow for testing the process by which the individual develops a sense of discomfort. So, it is necessary to work out special methods for assessing the severity of simulator sickness which would avoid these shortcomings. These methods should include both objective and subjective measures, which both reflect the person's interaction with virtual reality systems. Attempts to use vegetative reactions as objective measures were made but were unsuccessful, because these reactions have been considered to reflect only unspecified stress states under the influence of the virtual environment (Harm, Schlegel, 2012).

Eye movement characteristics were also considered to be objective measures of the person's behavior during the observation of virtual events. Oculomotor activity has long been studied in the context of explaining simulator sickness. It was proposed that proprioceptive signals of the eye muscles were one of the reasons for the occurrence of the vection illusion where you had a moving virtual environment and a stationary observer (Ebenholtz, Cohen & Linder, 1994). Eye movements were considered to be indicators of vestibular dysfunction which leads to the appearance of the vection illusion. Moreover, it was suggested that the vestibular-ocular reflex plays an important role in the appearance of the observer's illusory movements in the virtual environment (Smit, 2005; Authie & Mestre, 2011).

Method

The aim of the present research is to study oculomotor characteristics as indicators of a person's ability to resist motion sickness symptoms. The problem of the impact of virtual reality systems on the person's behavior mains poorly understood despite its high significance. We hypothesize that people with a vestibular system developed as a result of their professional activity, are less susceptible to simulator sickness compared with people who have no such professional abilities. The differences in a person's ability to resist the effects of the virtual environment may be revealed in the oculomotor characteristics registered during his/her interaction with virtual environment.

Participants

The ability to resist motion sickness induced by virtual reality devices was tested in experimental (90 observers: 29 females and 61 males) and control (20 observers: 11 females and 9 males) groups. The experimental group consisted of professional athletes in three sports — figure skating, wushu fighting, and football. In order to assess how athletes adapt to the virtual reality environment, we tested 30 figure skaters (18 females and 12 males, age range 15–24) with a high level of professional skill–21 figure skaters had a "master of sport" rank and 9 "candidate for master of sport" rank; 30 football players (30 males, age range 15–20) also with a high level of professional skill–7 football players had "master of sport" and 23 "candidate for master of sport" rank; and 30 wushu fighters (11 females and 19 males, age range 16–21) with a high level of professional skill — 20 of them had "master of sport" rank and 10 "candidate for master of sport" rank. As a control group, we used 20 students from Lomonosov MSU (9 males and 11 females, age range 18–24), who were not professional athletes. All participants had normal or corrected-to-normal vision and had no organic lesions of the vestibular system, or brain injury.

The reason for choosing the above-mentioned athletes was the following. Figure skaters, wushu fighters, and football players are considered highly professional elite athletes (Hutter, Oldenhof-Veldman & Oudejans, 2015). It was shown that these athletes had been trained since their childhood to perform a variety of complex elements (rotations, reversals, etc.), so they had well-developed vestibular systems. Typical exercises in figure skating include many elements with mostly acyclic movements. The large variety of movement forms involved lead to the development of different types of locomotor coordination (Absalyamova, Belyaeva & Zhgun, (1992). Mastering this sport has a great impact not only on the locomotor system but also on the sensory system functions (Mishin, 1985).

For wushu fighters the most important task is to sustain a posture and coordinate precise movements. Like figure skaters, wushu fighters have to execute a program consisting of multiple consecutive locomotion elements. Pre-defined movement rhythm and limited space, along with the large variety of complex elements, require a highly developed vestibular function. However, the professional activities of wushu fighters differ from the activities of figure skaters with respect to the need for movement speed and accuracy (Volkov, 2002).

Another sport which involves a high degree of dynamic abilities is football. Football players should be able to control the game situation on a large area of a football field, so central and periphery visual signals are an important part of their vestibular system development. These activities are similar to figure-skating activities, where athletes need to control their central and peripheral visual perspective, and resist the narrowing of the field of vision during the execution of a specific element (McLeod et al., 2008).

According to our initial hypothesis, the figure skaters were expected to demonstrate a higher ability to resist motion sickness symptoms than the other groups of athletes. Of the latter, wushu fighters were thought to be more successful in resisting motion sickness than football players. These assumptions were formulated on the basis of an analysis of athletes' professional activity. Their specific characteristics (main goals, means, and desired results) are presented in Table 1.

Table 1. Athletes' activity during competition

	Activity goals	Activity means	Desired results
Figure skaters	Artistry Program complexity Movement precision Movement elements Accuracy	Artistic image Vestibular function Coordination Balance Reaction speed	Demonstration of professional skills during competition Precise execution of the elements Victory in the competition
Wushu fighters	Movement precision Movement elements Accuracy	Vestibular function Coordination Balance Reaction speed	Demonstration of professional skills during competition Precise execution of the elements Victory in the competition
Football players	Execute movement sequences, leading to the team's victory, directly and indirectly	Coordination Vestibular function Reaction speed Visual control of wide field of vision	It depends on the athlete's role: for the goalkeeper it is the number of goals conceded; for the forward, the number of goals scored

It is necessary to note that our analysis was focused on the athlete's performance during competition, since that's when his/her professional skills are manifest most clearly. As seen in *Table 1*, the number of professional tasks for football players is limited. This may be understood as due to the fact that a narrower range of skills is required if one is training specifically to become a professional goalkeeper, a full-back, or a forward. On the contrary, wushu fighters and figure skaters have to possess a much greater range of skills, including artistry and mastering the complexity of the program.

Figure skating is the hardest sport in this regard, because it combines these requirements with a need for high speed of execution and precise coordination of athlete's body parts. So a well-developed vestibular system may be considered the most important factor for effective execution of the necessary elements — rotations, acrobatic elements, jumping, etc. Since the vestibular system is such a significant quality for a professional figure skater, it should be developed during very early professional specialization (at about 4–5 years of age), and should be tested throughout the period of professional growth. Wushu fighters possess a similar set of skills; however, due to the smaller percentage of complex rotations required, the role of the "vestibular" factor may be not as significant for their professional training.

Stimuli

To simulate motion sickness symptoms, special virtual environments were constructed. They consisted of a set of bright blue balls (total = 256), which spun as a whole around the observer. The average speed of the balls' rotation was 24°/s. To vary the impact of the virtual reality environments, three different virtual scenes were constructed which differed in the width of their visual fields. It was shown that the wider the visual field the observer used, the stronger was the simulator sickness (Menshikova, Kovalev, Klimova, Chernorizov, & Leonov, 2014). Thus, changes in the width of the viewing angle could gradually alter the severity of simulator sickness symptoms. Virtual scenes with three conditions were designed: in the first condition, all balls subtended 45° (the central part of the frontal screen); in the second, the balls subtended 90° (the whole frontal screen); in the third, they subtended 180° (the frontal and two side screens). The balls' diameters were 0.1°, 0.3° and 0.5° in the first, second, and third conditions, respectively.

Apparatus

The CAVE virtual reality system was used for the stimulation presentation (Fig. 1). The device had four large flat screens (Barco ISpace 4), which were connected into one cube consisting of three walls and a floor. The length of screen side was about 2.5 meters. Shutter eye glasses CrystalEyes 3 Stereographics were used. The projection system was based on BarcoReality 909. The projector's matrix resolution was 1400x1050 with 100 Hz update frequency. The tracking system was produced by ArtTrack2. VirTools 4.0 was used for software development. It supported DX9/GL2, HAVOK, particle systems, and shaders. The laboratory room was darkened. There were no any light sources except the CAVE systems projectors. The luminance range in stimulus scene was 1:230. The maximum luminance was 5.5 cd/m^2, the minimum 0.02 cd/m^2.

Eye movements were registered with SMI Eyetracking glasses 2.0, which have a resolution of 0.5° and a refresh rate of 60 Hz.

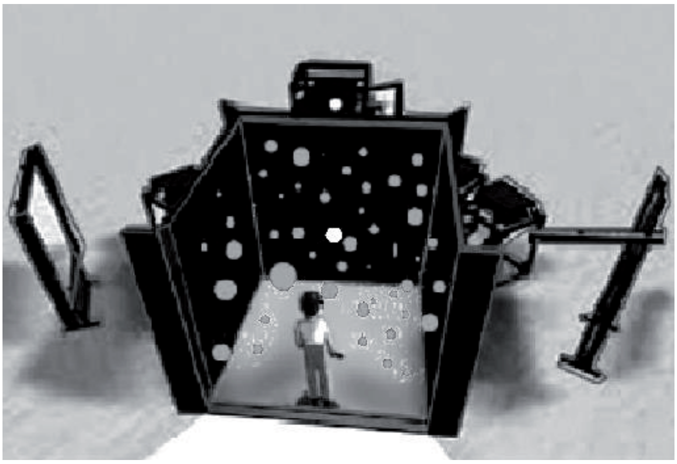


Figure 1. The main view of the CAVE system

Procedure and plan

The participant stood motionless in front of the central screen at a distance of 2.5 m observing the rotation of virtual balls around her/him. The rotation was carried out along the curved trajectory and lasted for two minutes for each condition. There was a fixation point in the center of the frontal screen (size: 0.5°; red in color). The whole experiment consisted of three presentations differed in the width of visual field (three conditions). The order of presentations was randomized. At the end of each presentation the participant completed the modified version of the SSQ questionnaire and then had a rest of 5 minutes. The entire experiment lasted approximately 25 minutes. At the end of the experiment, the participant was asked to report the feelings he experienced during interaction with the virtual environment.

Measurements

Objective and subjective measurements were used to evaluate the severity of the participants' motion sickness . Different eye movement characteristics (number of fixations, blinks, and saccades, saccade amplitudes) were used as objective measures, along with the ranking of the SSQ questionnaires as subjective measures. The data obtained was analyzed with SPSS 21.

Results

The data was analyzed using a two-way analysis of variance (ANOVA). The multidimensional tests showed statistically significant impacts of the factors "Number of conditions" ($F(16) = 77.86$, $p = 0.001$) and "Affiliation to a professional group" ($F(24) = 2.54$, $p = 0.001$) on the variables. Furthermore, the interaction between factors also reached a level of significance ($F(48) = 2.47$, $p = 0.001$). Thus, we showed that levels of oculomotor characteristics and questionnaire scores vary according to "professional group" affiliation. All conditions differed significantly according to oculomotor characteristics and to total questionnaire scores, as was revealed by the intergroup contrasts method: "blink count" ($F(3)=6.62$, $p<0.01$), "fixation count" ($F(3)=4.15$, $p<0.01$), "saccade count" ($F(3)=6.05$, $p<0.01$), "saccade amplitude" ($F(3)=16.39$, $p<0.01$), "Total score" ($F(3)=18.18$, $p<0.01$).

The data on the SSQ questionnaires was summarized separately for the various conditions and professional groups. The mean values of SSQ total scores for the experimental (figure skaters, wushu fighters, football players) and control groups are shown in *Fig. 2* for the first, second, and third viewing conditions (45^o = black columns, 90^o = gray columns, and 180^o = striped columns). According to the schedule, SSQ score values significantly differ ($F(1)=57.36$, $p<0.001$) between conditions for any group except the skaters group. For this group the increase of the field of vision (from 90^o up to 180^o) was not accompanied by an increase in total score. This result agrees with the skaters' self-reports: they felt no symptoms of strong discomfort during the observation of any moving virtual environments. On the contrary, all the other groups (wushu fighters, football players, and the control group) had twice increased their SSQ Total scores (up to 800 points) when the field of vision was changed from 90^o to 180^o. In self-reports they described their feelings as "dizziness with open eyes" and "nausea," and also noted increased difficulties in maintaining their gaze. Some participants complained they could hardly wait until the end of the experiment.

Data on oculomotor characteristics (blink, fixation, and saccade counts; saccade amplitudes) were also analyzed separately for the three conditions and professional groups. It was shown that all viewing conditions differed significantly ($F(16)=77.87$, $p<0.001$) according to the mentioned eye movement characteristics. The mean number of blinks, fixations, and saccades per minute for all three conditions (first = 45^o, second = $90^{o,}$ and third = 180^o, respectively) and for all groups are shown in *Fig. 3.* Their analysis shows that the mean values of oculomotor parameters differ insignificantly between the subject groups for the first ($t(29)=0.81$,

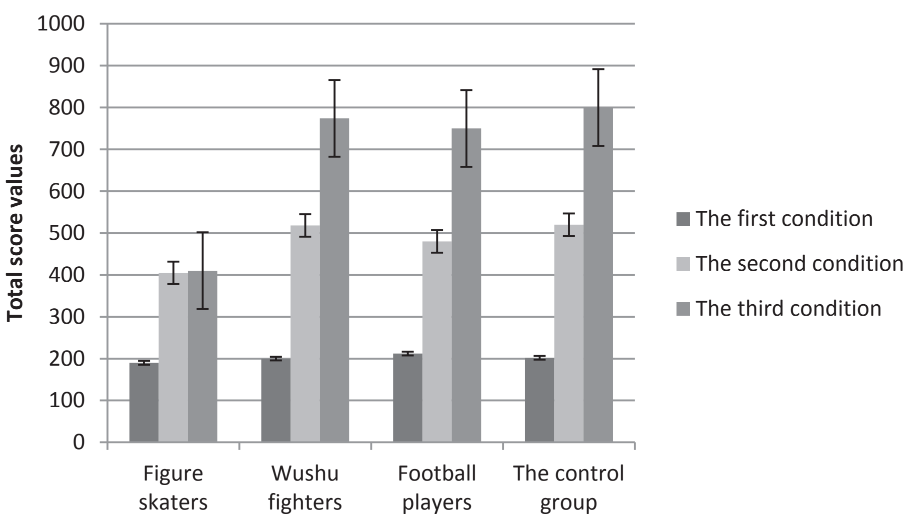


Figure 2. Mean values of SSQ Total scores for the experimental (figure skaters, wushu fighters, football players) and control groups in the first, second, and third viewing conditions

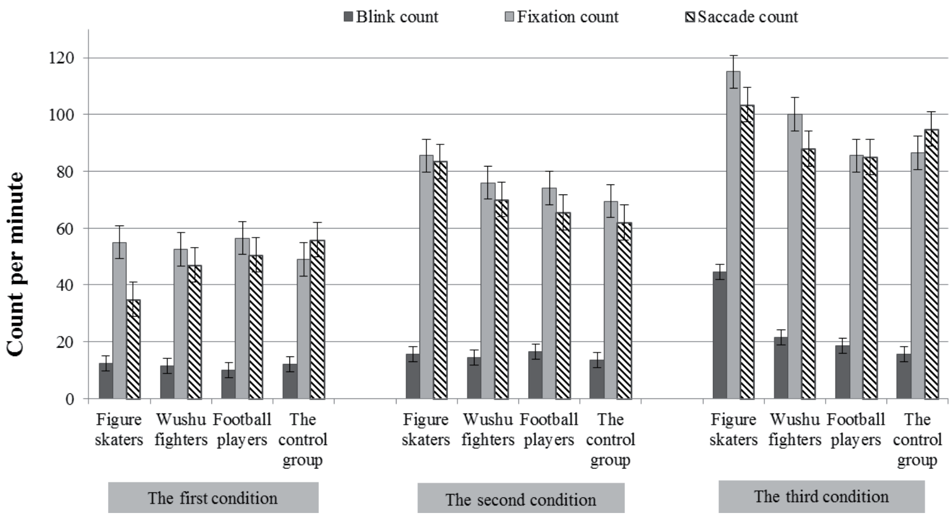


Figure 3. Mean values of blink, fixation, and saccade counts for the experimental (figure skaters, wushu fighters, football players) and control groups in the first, second, and third viewing conditions

$p>0.05$) and second conditions ($t(29) = 0.77$, $p < 0.05$). However, the third condition is significantly differentiated from the others in respect to oculomotor parameters: blink, fixation, and saccade counts increase in comparison with the first and second conditions. In particular, figure skaters have significantly greater oculomotor characteristics for the third condition ($t(29) = 3.34$, $p < 0.01$). This result is consistent with the SSQ Total scores. Mean values of saccade amplitudes for all three conditions (first = 45°, second = 90° and third = 180°, respectively) and for all groups are shown in Figure 4. The analysis of oculomotor characteristics of the experimental and control groups has been done for each condition separately.

In the first condition, both athletes and participants from the control group have approximately the same number of blinks (23), fixations (106), and saccades (94) counts per minute. The saccade amplitudes also vary only slightly, and their average value is about 7.3°. The SSQ total score is not very high either: its average value is about 201 for all participants (Fig. 2). This data is consistent with the self-reports by all participants which we received after their observation of the moving virtual environment in the first condition: they did not mention any feelings of discomfort.

The second condition differs from the first in that there was a substantial increase in all oculomotor characteristics in all subject groups. However, it should be noted that for figure skaters, changes in eye movements differ from eye movement changes in other participants. For example, their saccade amplitudes increase on average by 12% whereas for other participants it increases by 40 %. As for fixation, blink, and saccade counts, their values increase more in the second condition than in the first. Specific changes in the skaters' eye movements may be explained by their ability to resist the moving environment due to their professional training. Furthermore, subjective levels of discomfort evaluated on the basis of the self-re-

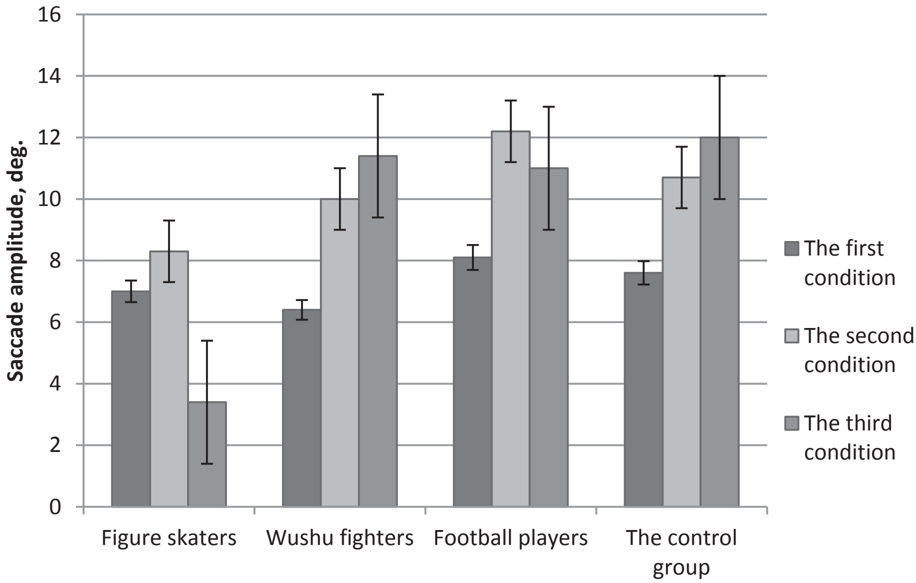


Figure 4. Mean values of saccade amplitudes for the experimental (figure skaters, wushu fighters, football players) and the control group in the first, second and third viewing conditions

ports and questionnaire scores were lower for the figure skaters than in the other participants.

Analysis of the third condition showed the highest results in almost all oculomotor characteristics for figure skaters. It should be noted that athletes of other groups (wushu fighters, football players) did not differ from untrained participants of the control group in respect to their oculomotor characteristics and SSQ scores.

Discussion

A figure skater's ability to successfully resist simulator sickness may be explained by his/her better developed vestibular function, which is considered the most important professional quality for this sport. Certain changes in oculomotor characteristics, in our opinion, are related to special aspects of figure skaters' athletic activity, which are actively developed from the early stages of professional specialization. Multiple accelerations and slowdowns, bows and rotations, and mastering the difficulties of maintaining balance on a limited supporting space, actively increase the ability of the body to evaluate its position and locomotion precisely, and thus lead to the development of the skater's vestibular function. At the same time many other functions are being developed (i.e., the so-called "feel of the ice")–muscular, joint and tactile sensitivities, eye estimation precision, differentiation of acoustic sensations, and the ability to integrate signals of different sensory systems during the execution of complex program elements (Chaikovskaya, 2003).

A number of studies have been devoted to highlighting vestibular habituation in populations who perform such complex program elements. In particular, they showed changes in the slow phase of vestibular ocular reflexes (VOR). In ballet dancers (Tschiassny, 1957) the VOR slow phase velocity is lower in comparison to non-trained participants. (Osterhammel et al. 1968). The VOR of gymnasts is characterized by a 15% shortening of slow phase and by a 25% decrease in saccadic amplitudes (Quarck et al., 1998). Among fighter pilots, the duration of the post-rotational nystagmus is also shorter (Aschan, 1954) than among civil aviation pilots.

We suggest that the reason for the differences in the characteristic eye movements of the figure skaters in our experiment is a change in their optokinetic nystagmus (OKN). OKN is similar to VOR in that OKN characteristics, such as a decrease in slow phase, may also be linked with habituation to their vestibular load. The fact that figure skaters have the lowest SSQ scores coheres with the habituation hypothesis since vestibular habituation seems to be accompanied by reduced motion sickness (MS). Thus, after a month of regular navigation, candidates for future maritime service become less sensitive to seasickness and show VOR habituation (Shupak et al. 1990). Repeated stimulation of the vestibular system as part of cosmonaut vestibular training also induced a decrease in MS (Clement et al. 2001). Repetitive vestibular stimulation can therefore cause changes in VOR and OKN, and at the same time a reduction in sensitivity to MS in virtual reality.

The analysis of the data on the football players revealed that negative symptoms (vertigo and nausea) appeared for the third condition (visual field of 180°). Along with the emergence of these symptoms, there was a lack of change in fixation and blink counts when compared with the first, second, and third conditions, which may be closely associated with their professional ability to control a whole visual field during the game. According to other authors, the ability to pay attention to objects located at the periphery of the visual field should be considered as one of the main professional qualities of football players (Williams, 2002). It was shown (Vaeyens, et al., 2007) that when the number of players in the environment increases, players change from exhibiting a low visual search rate with prolonged fixations, to a higher visual search rate of shorter duration. Evidently football players have started to realize their professional gaze behavior, which leads to an increase in the presence effect, and in the end to the higher vection strength. This assumption is in agreement with findings of a positive correlation between vection strength, motion sickness, and the presence effect.

Conclusion

The effectiveness of virtual reality systems for testing the professional abilities (namely the resistance to motion sickness) was demonstrated. These technologies allow researchers to initiate different kinds of vestibular function disturbances and to assess their strength in a real-time mode. Their application is effective for sportspeople of any age, especially for young athletes, who need to qualify to be selected for professional status in the early stages of their education.

A method based on eye movement characteristics can be successfully applied for testing resistance to motion sickness. Typical changes in eye movement characteristics were revealed in athletes, especially figure skaters, compared with participants not involved in professional sports. Athletes showed more active eye movements—an increase in blink, fixation, and saccade counts. The decrease in saccadic amplitudes was revealed for figure skaters. The eye movement data were consistent with Simulator Sickness Questionnaire scores.

Further research will be concentrated on the precise extraction gaze pattern, which is linked to the vection illusion, which can predict motion sickness. It is necessary likewise to clarify what exactly are the relevant stimulus parameters that increase vection. Therefore the future virtual scenes will be different in content from naturalistic visual stimuli to optokinetic drums to estimate the high-level and low-level influences on resistance to motion sickness.

Finally, in this study we did not investigate possible multi-sensory influences on motion sickness. The question we will ask in future experiments is how much visually-induced motion sickness in athletes is increased by adding other sensory modalities.

Acknowledgments

This work was supported by the Russian Foundation for Humanities (Grant no. 16-06-00312) and by the M.V. Lomonosov Moscow State University Program of Development.

References

Absalyamova, I.V., Belyaeva, A.Yu., & Zhgun, E.V. (1992). *Posobie po sinkhronnomu kataniyu na konkakh: Tochnye linii* [The handbook of synchronized skating: Accurate lines]. Moscow: GTsOLIFK.

Aschan, G. (1954). Response to rotatory stimuli in fighter pilots. *Acta Otolaryngologica, 116*, 24–31. doi: 10.3109/00016485409130269

Authié, C.N., & Mestre , D.R. (2011). Optokinetic nystagmus is elicited by curvilinear optic flow during high speed curve driving. *Vision Research*, *16*, 1791–1800. doi: 10.1016/j.visres.2011.06.010

Bailey, L., Denis, J.H., Goldsmith, G., Hall, P.L., & Sherwood, J.D. (1994). A wellbore simulator for mud-shale interaction studies. *Journal of Petroleum Science and Engineering, 11*, 195–211. doi: 10.1016/0920-4105(94)90040-X

Biocca, F. (1992). Will simulation sickness slow down the diffusion of Virtual Environment technology? *Presence: Teleoperators Virtual Environments, 1*(3), 334–343. doi: 10.1162/pres.1992.1.3.334

Brooks, J., Goodenough, R., Crisler, M., Klein, N., Alley R., Koon, B., … & Wills, R. (2010). Simulator sickness during driving simulation studies. *Accident Analysis & Prevention*, *42*(3), 788–796. doi: 10.1016/j.aap.2009.04.013

Chaikovskaya, E.A. (2003). *Figurnoe katanie* [Figure skating] Moscow: Fizkultura i sport.

Clement, G., Deguine, O., Parant, M., Costes-Salon, M., Vasseur-Clausen, P., & Pavy-LeTraon, A. (2001). Effects of cosmonaut vestibular training on vestibular function prior to spaceflight. *European Journal of Applied Physiology*, *85*, 539–545. doi: 10.1007/s004210100494

Ebenholtz, S.M., Cohen, M.M., & Linder, B.J. (1994). The possible role of nystagmus in motion sickness: A hypothesis. *Aviation Space and Environmental Medicine, 65,* 1032–1035.

Griffin, M. (2012) *Handbook of human vibration*. London: Academics Press.

Harm, D.L., & Schlegel, T.T. (2012). Predicting motion sickness during parabolic flight. *Autonomic Neuroscience: Basic and Clinical*, *97*(2), 116–121. doi: 10.1016/S1566-0702(02)00043-7

Hettinger, L.J., Berbaum, K.S., Kennedy, R.S., Dunlap, W.P., & Nolan, M.D. (1990). Vection and simulator sickness. *Military Psychology*, *2*(3), 171–181. doi: 10.1207/s15327876mp0203_4

Hettinger, L.J. (2002). Illusory self-motion in virtual environments. In K.M. Stanney (Ed.), *Handbook of virtual environments* (pp. 471–492). Mahwah, NJ: Lawrence Erlbaum.

Howarth, P.A., & Costello P.J. (1997). The occurrence of virtual simulation sickness symptoms when an HMD was used as a personal viewing system. *Displays, 18*(2), 107–116. doi: 10.1016/S0141-9382(97)00011-5

Hutter, R.V., Oldenhof-Veldman, T.M., & Oudejans, R.R. (2015). What trainee sport psychologists want to learn in supervision. *Psychology of Sport and Exercise, 16*, 101–109. doi: 10.1016/j.psychsport.2014.08.003

Jing, C.D.H. (1998). Competitive series of skills and tricks in wushu – the main way to preserve and develop Chinese wushu. *Journal of Chengdu Physical Education Institute*, *1,* 372–384.

Kellogg, R.S., Kennedy R.S., & Graybiel, A. (1964). Motion sickness symptomatology of labyrinthine defective and normal subjects during zero gravity maneuvers. *Aerospace Medicine*, *36*, 315–318.

Kennedy, R.S., Hettinger, L.J., & Lilienthal, M.G. (1990). Simulator sickness. *Motion and Space Sickness*, *1*, 317–341. . doi: 10.1207/s15327108ijap0303_3

Kennedy, R.S., Lane, N.E., Kevin, S., Berbaum, K.S., & Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, *3*(3), 203–220.

Keshavarz, B., & Berti, S. (2014). Integration of sensory information precedes the sensation of vection: A combined behavioral and event-related brain potential (ERP) study. *Behavioural Brain Research*, *259*(1), 131–136. doi: 10.1016/j.bbr.2013.10.045

McCauley, M.E., & Sharkey, T.J. (1998). Cybersickness: Perception of motion in virtual environments. *Presence: Teleoperators Virtual Environments, 1*(3), 311–318. doi: 10.1162/pres.1992.1.3.311

McLeod, P., Reed, N., Gilson, S., & Glennerster, A. (2008). How soccer players head the ball: A test of optic acceleration cancellation theory with virtual reality. *Vision Research*, *48*(13), 1479–1487. doi: 10.1016/j.visres.2008.03.016

Menshikova, G., Kovalev, A., Klimova, O., Chernorizov, A., & Leonov, S., (2014). Testing the vestibular function development in junior figure skaters using the eye tracking technique. *Procedia – Social and Behavioral Sciences*, *146*(25), 252–258. doi: 10.1016/j.sbspro.2014.08.123

Mishin, A. N. (1985). *Figurnoe katanie na konkakh: uchebnik dlya institutov fiz kultury* [Figure skating on ice: The handbook for sport institutes]. Moscow: Fizkul'tura i sport.

Mon-Williams, M., Wann, J.P., & Rushton, S. (1995). Design factors in stereoscopic virtual-reality displays. *Journal of the Society for information Display*, *3*(4), 207–210. doi: 10.1889/1.1984970

Osterhammel, P., Terkildsen, K., Zirstorf, K. (1968). Vestibular habituation in ballet dancers. *Acta Otolaryngologica, 66*, 221–228. doi: 10.3109/ 00016486809126289

Quarck, G., Etard, O., Darlot, C., & Denise, P. (1998). Motion sickness susceptibility correlates with otolith- and canal-ocular reflexes. *Neuroreport*, *9*, 2253–2256. doi:10.1097/00001756-199807130-00019

Reason, J.T. (1978). Motion sickness adaptation: A neural mismatch model. *Journal of the Royal Society of Medicine*, *71*, 819–829.

Riccio, G.E., & Stoffregen, T.A. (1991). An ecological theory of motion sickness and postural instability. *Ecological Psychology*, *3*(3), 195–240. doi: 10.1207/s15326969eco0303_2

Sharples, S., Cobb, S., Moody, A., & Wilson, J. R. (2008). Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays*, *29*(2), 58-69. doi: 10.1016/j.displa.2007.09.005

Shupak, A., Kerem, D., Gordon, C., Spitzer, O., Mendelowitz, N., & Melamed, Y. (1990). Vestibulo-ocular reflex as a parameter of seasickness susceptibility. *Annals of Otology, Rhinology, and Laryngology, 99*, 131–136. doi: 10.1177/000348949009900211

Smit, K., & Yu, M. (2005) *Biologiya sensornykh sistem* [Biology of sensory systems]. Moscow: Binom.

Stanney, K.M., Hale, K.S., Nahmens, I., & Kennedy, R.S. (2003). What to expect from immersive virtual environment exposure: Influence of gender, body mass index, and past experience. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *45*(3), 504–520. doi: 10.1518/hfes.45.3.504.27254

Timmis, A., Turner, K., & van Paridon, K. (2014). Visual search strategies of soccer players executing a power vs. placement penalty kick. *PLoS ONE, 9*(12), e115179. doi:10.1371/journal.pone.0115179

Tschiassny, K. (1957). Studies concerning vestibular factors in the ballet dancer, the pigeon and the blind person. *Transactions. Section on Otolaryngology American Academy of Ophthalmology and Otolaryngology journal, 61*(4), 503–506.

Vaeyens, R., Lenoir, M., Philippaerts, M., & Williams, A.M. (2007). Mechanisms underpinning successful decision making in skilled youthful soccer players: An analysis of visual search behaviors. *Journal of Motor Behavior, 39*, 395–408. doi: 10.3200/JMBR.39.5.395-408

Volkov, I.P. (2002). *Sportivnaya psikhologiya v trudakh otechestvennykh spetsialistov* [Sport psychology in domestic specialists review]. Moscow: Piter.

Williams, A.M. (2000). Perceptual skill in soccer: Implications for talent identification and development. *Journal of Sport Sciences*, *18*, 1–14. doi: 10.1080/02640410050120113

Zinchenko, Yu.P., Menshikova, G.Ya., Bayakovsky, Yu.M., Chernorizov, A.M., & Voiskounsky, A.E. (2010). Technologies of virtual reality in the context of World-wide and Russian psychology: methodology, comparison with traditional methods, achievements and perspectives. *Psychology in Russia. State of the Art, 3,* 12–45. doi: 10.11621/pir.2010.0001

Original manuscript received April 29, 2016
Revised manuscript accepted October 10, 2016
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society | Lomonosov Moscow State University

The relationship between interference control and sense of presence in virtual environments

Boris B. Velichkovsky

Faculty of Psychology, Lomonosov Moscow State University, Moscow, Russia

Corresponding author. E-mail: velitchk@mail.ru

Background. The sense of presence is an important aspect of interaction with virtual reality applications. Earlier we suggested that presence can depend on cognitive control. The latter is a set of meta-cognitive processes which are responsible for configuring the cognitive system for the accomplishment of specific tasks with respect to a given context. In particular, cognitive control helps in preventing interference from the task-irrelevant variables.

Objective. This study aimed at investigation of the possible relationship between interference control and aspects of presence.

Design. Thirty-nine subjects (32 female and 7 male, aged 18 to 27 years) participated in the study. The subjects were assessed via a battery of interference control tasks (Flanker Task, Go/No Go task, antisaccade task) and performed a virtual scenario (navigating within an array of randomly placed virtual digits in correct numerical order) in high-immersion (CAVE) and low-immersion (standard computer display) virtual environments. Afterwards, the subjects completed a Russian version of the ITC-Sense of Presence inventory.

Results. We found that interference control is generally related to the sense of presence, especially in the CAVE (high-immersion) environment. Sensory interference control was most strongly associated with various aspects of presence (overall presence score, spatial presence, and emotional involvement). Motor interference control was associated with spatial presence and emotional involvement, but this relationship was weaker than was the case with sensory interference control. Low-immersion virtual environments attenuate some of these links between interference control and presence so that only sensory interference control remains a notable predictor of presence. **Conclusion.** Interference control is positively associated with presence in virtual environments with varying immersion levels. This may reflect a more general cause-and-effect relationship between cognitive control and the feeling of presence in virtual reality.

Keywords: virtual reality, presence, interference, cognitive control, attention, flanker task, antisaccade task, Go/No Go task

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0311
http://psychologyinrussia.com

Introduction

Virtual reality (VR) applications are being increasingly used in science and technology. An important aspect of interacting with VR is presence (Sanchez-Vives & Slater, 2005; Diemer et al., 2015). Presence is associated with the feeling of being located in the virtual environment and the realness of interacting with virtual objects. It can be defined as the "perceptual illusion of non-mediation" (Lombard & Ditton, 1997). This means that presence is experienced as if the virtual environment is real and there is no or little conscious awareness of either the real environment, or of the technology used to produce the virtual experience. It is obvious that presence is the very heart of the idea of VR; if there is no presence, the user experiences discomfort when interacting with a VR application, and there is little opportunity for virtual reality to become the user's reality. The feeling of presence exerts a strong influence on the quality of interaction with VR systems. It is easy to see that optimizing presence is an important objective for designers of VR applications.

There are many determinants of presence, both technological and psychological. Technological determinants of presence mostly concern the fidelity of the VR presentation. The rule of thumb is that the more realistic the output of the presentation system is the more presence emerges in the user. Thus, full immersion (via head-mounted displays or the CAVE system), high frame rate, high resolution, the use of spatial sound, low feedback latency, the possibility of interacting with virtual objects should produce increased presence. However, this is not always the case. An example is the so-called "book paradox": a well-written book produce high presence and involvement in the reader although there is no immersion at all. Presence is a subjective phenomenon and therefore psychological factors are at least as important as the visual fidelity of the presentation.

Research on the psychological determinants of presence has produced some interesting results. Presence seems to be dependent upon gender, age, and personality variables (Sacau et al., 2008). Among the latter, extraversion, locus of control, openness to experience, and psychological absorption were shown to be of relevance (Sas, 2004; Baños et al., 1999). However, the results have been mixed. For instance, extraversion was shown to be related to presence both positively and negatively. These mixed results indicate that personality traits may not be an appropriate level of organization on which to search for psychological determinants of presence. Previously, it has been suggested that presence can depend on cognitive control (Velichkovsky, 2014). Cognitive control is a set of meta-cognitive processes which are responsible for configuring the cognitive system for the accomplishment of specific tasks with respect to a given context (Notebaert & Verguts, 2008). As such, cognitive control may be crucially responsible for re-configuring the cognitive system towards optimal interaction with a VR environment which is different from interaction with the real environment. In this study, we investigate whether presence is related to a fundamental aspect of cognitive control – specifically, control of interference.

Control of interference (or inhibition) is considered a basic function of cognitive control (Miyake et al., 2000). It comprises a set of related functions (Friedman & Miyake, 2004). These include control of sensory interference (interference produced by the presence of irrelevant sensory stimuli); control of cognitive interference (control of irrelevant representations activated in the course of cognitive processing and competing for processing resources, including proactive and reactive interference phenomena in memory); and control of inappropriate response tendencies. Inhibition of inappropriate saccadic responses (as measured by the antisaccade task) is a special class of inhibition functions, related but not identical to the inhibition of inappropriate motor responses. These functions are functions of voluntary interference control to be differentiated from automatic inhibition phenomena like inhibition-of-return. Research shows that inhibition functions are at the core of the more complex executive functions responsible for voluntary control of behavior (Friedman & Miyake, 2004). The age-related decrease in the efficiency of interference control may be the cause of cognitive decline observed among older persons (Hasher & Zacks, 1988).

There are several paths through which interference control may be related to presence. First, **control of sensory interference** may help the user devote attention to the virtual environment and ignore distracting stimuli from the now irrelevant real environment. This is the more important as the investment of attentional resources is considered to be an objective indicator of presence (Draper et al., 1998). Second, **control of cognitive interference** helps the mind free central resources from processing thoughts which are irrelevant to the virtual environment and concentrate on building an appropriate mental model of interacting with the virtual environment. Third, **efficient control of eye movements** would assist the user in directing his/her overt attention toward the virtual environment with the aim of preferentially processing VR stimuli. Fourth, **control of motor responses** would help the user to selectively attend to responses appropriate to the given virtual stimulation. In sum, interference control may assist the user in processing the stimuli, representations, and responses relevant to interacting with the virtual environment and thus help him/her ignore irrelevant stimuli, representations, and responses which would otherwise hinder the interaction with the virtual environment.

In the present study, correlations between the efficiency of interference control and the subjective sense of presence are investigated. Several interference control tasks are used to assess different aspects of interference control. The aim of the study is to show that there is a positive relationship between the efficiency of interference control and the sense of presence. To the best of our knowledge, this is the first study aimed at demonstrating this relationship. The results of the study will help to elucidate the cognitive mechanisms involved in the emergence of the sense of presence while interacting with virtual environments. They can also have practical implications, since finding stable relationships between cognitive control variables and presence may help in designing more effective VR applications, and in the selection of users for interacting with virtual environments.

Method

Subjects

Thirty-nine subjects aged 18 to 27 years–32 female and 7 male–all students at M.V. Lomonossov-Moscow State University, took part in the study.

Interference control tasks: Flanker Task

The stimuli were five horizontally oriented arrows, arranged in a congruent (>>>>>, <<<<<) or an incongruent (>><>>, <<><<) order. The subject's task was to identify the direction of the middle arrow by pressing a key. The subject was given a training series with 36 trials and a main series with four blocks. In each block, every sequence of arrows was presented 36 times. The presentation time was 1500 milliseconds (ms), with an interstimulus interval of 1000 ms. Registered were the subject's reaction time (T_{av}) and accuracy (A_{av}) in general and for each trial type (T_{con}, A_{con}, T_{inc}, A_{inc}), and time (T_{int}) and accuracy (A_{int}) related interference index (computed as the difference in responses between incongruent and congruent trials).

Interference control tasks: Go/No Go task

The stimuli were a target stimulus (X, 80% of presentations) and distractor stimuli (А, Г, Е, И, К, Л, М, Н, П, Т, О, Э, Ю, Я). The stimuli were presented randomly in the center of the screen. The subject's task was to press a key if the target stimulus was presented. The presentation time was 300 ms with an interstimulus interval of 700 ms. There was a training series with 20 trials and a main series with 200 trials. Registered were reaction time (T_{av}), accuracy (A_{av}), the number of hits (N_{hit}) and false alarms (N_{fa}).

Interference control tasks: Antisaccade task

A fixation point was presented in the middle of the screen for a varying amount of time (1500 to 3500 ms, in increments of 250 ms). A visual distractor (a square with a side of 0.4°) was presented on the a randomly selected half of the screen for 200 ms. The presentation of the distractor was followed by the presentation of the target stimulus (an arrow pointing left or right) in the opposite half of the screen for a very short period of 100 ms; the stimulus was masked after presentation. The subject's task was to identify the orientation of the target stimulus by pressing a key: a procedure which implies suppressing the reflexive saccade towards the distractor. There was a training series of 16 trials and a main series of 96 trials. Registered were reaction times (T_{av}) and accuracy (A_{av}).

Presence questionnaire

The subjective feeling of presence was assessed with a Russian version of the ITC-Sense of Presence Inventory (Lessiter et al., 2001). This is a questionnaire of 44 items with subscales of Spatial Presence (SP), Naturalness (N), Emotional Involvement (EI), and Negative Effects (NE). Spatial Presence pertains to the illusion of being transferred to the virtual environment. Naturalness pertains to ease of understanding the virtual environment. Emotional Involvement pertains to the sense of engagement and its pleasantness while interacting with the virtual environment. Negative Effects pertains to vestibular disturbances which may emerge as a consequence of increased presence (nausea). An overall index of presence (Presence) was obtained by summing the scores for all items.

Virtual scenario

To assess the sense of presence we needed a task which required orientation and movement in a virtual environment and active interaction with virtual objects. We created a virtual task which consisted in navigating within an array of randomly placed in a rectangular virtual space with a side of 20 m digits in correct numerical order. This virtual scenario was presented either in a high-immersion CAVE system or by the means of low-immersion standard 19" computer display. For each type of presentation, there was a training session (digits from 1 to 5) and a main session (digits from 1 to 9).

Procedure

The subjects first performed the interference control tasks. Afterwards, they performed the virtual scenarios with the order of presentation type (CAVE, display) counterbalanced across subjects. After completion of each virtual scenario, the subjects completed the presence inventory.

Results

Presence scores

The two virtual environments differed as to the intensity of subjective presence in the predicted way. Subjective presence was significantly higher in the high-immersion CAVE condition. This applied not only to the overall index of presence, but also to all components of subjective presence. Descriptive statistics on the subjective sense of presence and its components and the results of statistical analysis are presented in Table 1.

Table 1. Means (standard deviations) for the components of presence in different virtual environments and the results of statistical comparison.

		Presence components			
	Presence	**SP**	**N**	**EI**	**NE**
CAVE	128.6 (17.9)	2.93 (.47)	3.35 (.61)	2.73 (.63)	2.15 (.72)
Display	96.1 (16.5)	2.04 (.45)	2.63 (.61)	2.00 (.49)	1.92 (.60)
Wilcoxon T-test (Z)	-5.65 $p<0.001$	-5.65 $p<0.001$	-5.15 $p<0.001$	-5.05 $p<0.001$	-2.52 $p<0.05$

Flanker Task

Non-parametric correlations between Flanker Task variables and components of presence are presented in *Table 2*. In the high-immersion CAVE condition, components of presence are related to average accuracy (overall presence, SP, EI, and NE components); accuracy in the incongruent condition (overall presence and NE component); and accuracy-related flanker interference cost (NE component).

Other correlations were not significant even at a liberal p<0.1 level. For the low-immersion display conditions, almost all correlations were not significant. Only average accuracy in Flanker Task was related to the overall presence score. Generally, correlations between the efficiency of the Flanker Task and the components of presence were lower in the low-immersion display condition than in the CAVE condition.

Table 2. Non-parametric correlations between the efficiency of the Flanker Task and components of presence (** = p<0.05, * = p<0.1).

	Presence components				
	Presence	**SP**	**N**	**EI**	**NE**
			CAVE		
T_{av}	–.028	–.047	.091	–.229	.016
A_{av}	*.519** *	*.455** *	.229	*.421** *	.256
T_{con}	.076	–.106	–.062	–.253	.016
T_{inc}	–.004	–.009	.087	–.240	–.055
A_{con}	–.231	–.173	–.207	–.162	.030
A_{inc}	*–.265**	–.258	.004	–.216	*–.291**
T_{int}	.106	.148	.035	–.021	.080
E_{int}	.213	.215	–.037	.182	*.281**
			Display		
T_{av}	–.005	.055	–.121	.127	.105
A_{av}	*.256**	–.144	–.158	–.108	–.235
T_{con}	.002	.069	–.130	.126	.122
T_{inc}	.043	.053	–.035	.155	.071
A_{con}	–.176	–.136	–.124	–.110	–.028
A_{inc}	–.202	–.149	–.071	–.108	–.210
T_{int}	.072	–.057	–.130	.007	–.071
E_{int}	.168	.123	.045	.082	.144

Go/No Go task

Non-parametric correlations between Go/No Go task variables and presence components are presented in Table 3. Under the high-immersion CAVE conditions, components of presence are related to average accuracy and the number of false alarms (EI component). In the low-immersion display condition, presence (SP component) is related to average reaction time. All other correlations were not significant even at a liberal p<0.1 level, and generally correlations between Go-No Go task efficiency and components of presence were lower than the correlations between Flanker Task efficiency and components of presence.

Table 3. Non-parametric correlations between the efficiency of Go/No Go task and components of presence (* = p<0.1)

	Presence components				
	Presence	**SP**	**N**	**EI**	**NE**
			CAVE		
T_{av}	.009	.173	–.089	–.133	.039
A_{av}	–.172	–.100	–.098	*.311**	–.140
N_{hit}	–.025	–.089	.004	–.098	.058
N_{fa}	.154	.070	.101	*–.297**	.139
			Display		
T_{av}	.138	*–.293**	–.033	.077	.055
A_{av}	–.222	–.132	–.231	–.213	.115
N_{hit}	–.048	–.036	.099	–.093	–.008
N_{fa}	.199	.100	.226	.192	.126

Antisaccade task

Non-parametric correlations between antisaccade task variables and components of presence are presented in *Table 4*. Under the high-immersion CAVE conditions the only correlation obtained was between SP component of presence and antisaccade task accuracy (this correlation was also increased in the display condition relative to most other correlations, but failed to reach the selected level of significance). Under the low-immersion display condition, there were no significant correlations between antisaccade task variables and components of presence. Overall, the correlations between the antisaccade task variables and components of presence were similar to that between the Go/No Go task variables and presence, and lower than that between the Flanker Task and presence.

Table 4. Non-parametric correlations between antisaccade task efficiency and components of presence (* = p<0.1).

	Presence components				
	Presence	**SP**	**N**	**EI**	**NE**
			CAVE		
T_{av}	–.010	.012	–.008	–.127	.205
A_{av}	–.048	*.261**	.196	–.164	.030
			Display		
T_{av}	–.006	.074	–.094	.126	.167
A_{av}	–.236	–.225	–.102	.019	–.113

Discussion

In this study, the relationship between interference control and the subjective sense of presence while interacting with virtual environments was assessed in a sample of university students. The aim of the study was to check whether interference control–as a part of cognitive control–is a reliable determinant of presence. Several interference control tasks were used to measure the efficiency of various aspects of interference control: the Flanker Task, the Go/No Go task, and the antisaccade task. For instance, the Flanker Task addressed the efficiency of sensory and cognitive interference control, the Go/No Go task addressed the efficiency of cognitive and motor interference control, and the antisaccade task captured the efficiency of inhibiting inappropriate saccadic eye movements and control of overt attention. Besides investigation of the general relationship between interference control and presence, the study raised the interesting research question of whether different aspects of interference control are differentially related to various presence components.

We also contrasted two virtual environments. The CAVE environment is characterized by highly intense immersion of the participants and thus should produce conditions which promotes emergence of presence. The standard display environment is characterized by low immersion and thus is less able to promote a strong sense of presence. We checked whether interference control is differentially related to presence in virtual environments with different level of immersion. It is reasonable to assume that under different levels of immersion, there are different cognitive mechanisms which lead to interference control influencing presence.

The study results showed that performance in the Flanker Task is strongly related to various components of presence in the CAVE environment. Accuracy in the Flanker Task–reflecting the effectiveness of suppressing irrelevant visual stimuli competing for processing with the focal target stimuli–was strongly related to the overall index of presence, spatial presence component, and the emotional involvement component. That is, there seems to be a strong generalized relationship between presence and the effectiveness of irrelevant stimuli suppression in a high-immersion virtual environment. This result is in perfect accord with the notion that effectively directing a person's attention toward virtual environment objectively determines presence (Draper et al., 1998). The ability to voluntarily direct attention toward relevant stimuli is captured by the Flanker Task, and this individual ability may be the reason why different people experience different levels of presence in the same virtual environment. Sensory interference control is thus a possible determinant of presence, at least in highly immersive virtual environments.

The relationship between sensory interference control and presence is less articulated in low-immersion virtual environments. Here, a generalized relationship between the accuracy of the Flanker Task and the overall index of presence can be found, but it is significant only on a tendency level. While overall the results obtained support the idea that sensory interference control is a reliable determinant of presence, low-immersion virtual environments exhibit factors which attenuate this relationship. One factor is, obviously, the abundance of irrelevant visual stimuli in VR user's field of view which compete for processing resources much stronger than is the case in high-immersion virtual environments. This makes sensory interference control and covert attention control (both are accessed by the Flanker Task) less effective means of filtering out irrelevant stimulation. The latter operation is now more effectively accomplished by voluntarily restricting saccadic eye movements that results in a narrowing the efficient field of view. A second conclusion which may be drawn from the data is that the Flanker Task may reflect primarily sensory interference control though cognitive interference control would also be important for producing a strong feeling of presence. As there is much more potential for activating irrelevant representations (that is, representations irrelevant to the VR experience) in low-immersion environments than in high-immersion ones, sensory interference control becomes a less effective predictor; the relationship between effectiveness of the Flanker Task and presence is attenuated in this case.

The Go/No Go task also was related to presence, although to a lesser extent than the Flanker Task. The Go/No Go task reflects the ability to suppress responses that are usually permissible but become inappropriate in a given context. It was found that for the high-immersive CAVE environment, the accuracy of the Go/No Go task and the number of false alarms were related to the emotional involvement component of presence. For the low-immersion display condition, the Go/No Go task was unrelated to the emotional involvement but only to the spatial presence component. The fact that motor inhibition is generally related to presence, at least in high-immersion environments, seems justified. Control of an inappropriate response means the re-direction of attention toward the execution of appropriate response. Thus, effective motor interference control helps to shape overt actions with respect to the actual context, which produces a seamless interaction with the current (virtual) environment. This, in turn, reduces the emotional discomfort the user may experience when interacting with an unfamiliar virtual environment, which promotes the feeling of emotional involvement. If the ability to suppress inappropriate responses is limited, on the other hand, then interaction with the unfamiliar virtual environment is hindered, and there is emotional discomfort which disturbs the feeling of emotional involvement and presence generally. That this relationship does not hold for low-immersion environments again shows that presence is produced by different cognitive mechanisms in high-immersion and low-immersion environments. For instance, it may be that emotional involvement is generally weak in low-immersion environments, and floor effects hinder the creation of a relationship between motor interference control and emotional involvement.

In low-immersion environment, however, there is a link between Go/No Go task reaction time and spatial presence. The reaction time measure reflects the efficiency with which motor interference control demands are processed. That this general control efficiency measure is related to spatial presence may be explained by efficient motor interference control promoting a quick re-organization of spatial perception and action control in respect to the spatial specifics of a given virtual environment. Thus, the relationship we obtained may reflect a more general relationship between presence and cognitive control (see below).

Performance in the antisaccade task was only weakly related to presence. There was a tendency for statistical significance of the relationship between antisaccade task accuracy and spatial presence in the high-immersion environment, and no relationship between antisaccade task performance and presence in the low-immersion environment. Spatial presence refers to the aspects of presence related to the illusion of being transferred into another (virtual) space and acting there as if it were the real space. An obvious factor relating antisaccade task and spatial presence is that effective control of saccades restricts the user's efficient field of view to the virtual environment only. However, the cognitive mechanisms involved in the antisaccade task and presence may be more complex. Antisaccade task performance is considered to be an index of attention control (Munoz & Everling, 2004) and the ability to maintain a goal in distractor-rich environments (Engle, 2002). Antisaccade task performance may thus be related to the general ability to perform goal-directed actions in an unfamiliar virtual environment in the presence of interference from the real environment. In this case, there is an overlap between mechanisms that produce the link between antisaccade task and presence, on the one hand, and between Flanker Task and presence, on the other. In this respect it is important to note that Flanker Task performance was also found to be related to spatial presence.

In general, our data suggest that interference control generally is related to presence, although this link is mostly pronounced for sensory interference control and for high-immersion virtual environments. These findings help explain why constructs only distantly related to presence like working memory capacity (WMC) have been shown to correlate with presence (Rawlinson et al., 2012). WMC has repeatedly been shown to be related to the efficiency of interference control (Engle, 2002), and the link between WMC and presence may be due to the involvement of interference control in both cognitive processes. As WMC is especially related to control of cognitive interference, this aspect of interference control–only partially represented in the present study–may be another important determinant of presence. It is a task for future research to show this link by unequivocally operationalizing cognitive interference control (for example, by the means of proactive interference control memory tasks).

The relationship we found between interference control and presence may also indicate that there is a connection between cognitive control (of which interference control is an important aspect) and the subjective feeling of presence in virtual environments. Cognitive control is conceptualized as the group of brain processes responsible for context-dependent cognitive system re-configuration toward the task at hand. As such, cognitive control is excellently suited for the task of tuning the cognitive system toward interaction with a virtual environment substituted for the real environment. Our study demonstrated how this theoretical conclusion can be experimentally elaborated by finding links between presence and diverse aspects of cognitive control. It is necessary to note that neurocognitive research also support the notion that there is a link between cognitive control and presence. For instance, Jäncke et al. (2009) have shown that prefrontal cortex activity associated with cognitive control may be associated with subjective feeling of presence.

Conclusion

In this study, the relationship between interference control and aspects of presence was investigated. It was found that interference control is generally related to the sense of presence. This link was especially strong for high-immersion virtual environments. From the various aspects of interference control, sensory interference control was most strongly associated with various aspects of presence (overall presence score, spatial presence, and emotional involvement). Sensory interference control exerts its influence on presence by allowing the user to concentrate on stimuli pertaining to the virtual environment, and to ignore irrelevant stimuli pertaining to the real environment. Motor interference control was associated with spatial presence and emotional involvement, but this relationship was less strong than was the case with sensory interference control. Motor interference control exerts its influence on presence by allowing the user to concentrate on VR appropriate actions, thus making the VR-user interaction more natural. Low-immersion virtual environments attenuate some of these links between interference control and presence so that only sensory interference control remains a notable predictor of presence in this case. The relationship we found between interference control and presence may reflect a more general relationship between cognitive control and presence.

Acknowledgments

This study is supported by the Russian Foundation for Basic Research, RFFR grant 15-06-08998.

References

Baños, R., Botella, C., Garcia-Palacios, A., Villa, H., Perpina, C., & Gallardo, M. (1999). Psychological variables and reality judgements in virtual environments: The role of absorption and dissociation. *Cyberpsychology and Behavior*, *2*, 135–142. doi: 10.1518/001872098779591386

Draper, J., Kaber, D., & Usher, J. (1998). Telepresence. *Human Factors*, *40*, 354–375. doi: 10.1518/001872098779591386

Diemer, J., Alpers, G., Peperkorn, H., Shiban, Y., & Mühlberger, A. (2015). The impact of perception and presence on emotional reactions: A review of research in virtual reality. *Frontiers in Psychology*, *6*, 26. doi: 10.3389/fpsyg.2015.00026

Engle, R. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*, 19–23. 10.1111/1467-8721.00160

Friedman, N., & Miyake, A. (2004). The relations among inhibition and interference control functions: a latent-variable analysis. *Journal of Experimental Psychology: General*, *133*, 101–135. doi: 10.1037/0096-3445.133.1.101

Hasher, L., & Zacks, R. (1988). Working memory, comprehension, and aging: A review and a new view. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation, Vol. 22* (pp. 193–225). New York: Academic Press. doi: 10.1016/S0079-7421(08)60041-9

Jäncke, L., Cheetham, M., & Baumgartner, T. (2009). Virtual reality and the role of the prefrontal cortex in adults and children. *Frontiers in Neuroscience*, *3*, 52–59. doi: 10.3389/neuro.01.006.2009

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments*, *10*, 282–297. doi: 10.1162/105474601300343612

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer Mediated Communication*, *3*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.1997.tb00072.x/abstract 10.1111/j.1083-6101.1997.tb00072.x

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. doi: 10.1006/cogp.1999.0734

Munoz, D., & Everling, S. (2004). Look away: The anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience, 5*, 218–228. doi: 10.1038/nrn1345

Notebaert, W, & Verguts, T. (2008). Cognitive control acts locally. *Cognition*, *106*, 1071–1080. doi: 10.1016/j.cognition.2007.04.011

Rawlinson, T., Lu, Sh., & Coleman, P. (2012). Individual differences in working memory capacity and presence in virtual environments. In H. Zhang, A. Hussain, D. Liu & Zhanshan Wang (Eds.), *Advances in brain inspired cognitive systems. Lecture Notes in Computer Science*, vol. 7366, (pp. 22–30). Heidelberg/New York: Springer. doi: 10.1007/978-3-642-31561-9_3

Sacau, A., Laarni, J., & Hartmann, T. (2008). Influence of individual factors on presence. *Computers in Human Behavior*, *24*, 2255–2273. doi: 10.1016/j.chb.2007.11.001

Sanchez-Vives, M., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, *6*, 332–339. doi: 10.1038/nrn1651

Sas, C. (2004). Individual differences in virtual environments. *Computational science – ICCS 2004, Fourth international conference, proceedings, part III. Lecture Notes in Computer Science*, vol. 3038, (pp. 1017–1024). Heidelberg/New York: Springer. doi: 10.1007/978-3-540-24688-6_131

Velichkovsky, B.B. (2014). Psychological factors of the emerging sense of presence in virtual environments. *National Psychological Journal*, *3*, 31–38. doi: 10.11621/npj.2014.0304

Original manuscript received August 22, 2016
Revised manuscript accepted January 25, 2017
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society
Lomonosov Moscow State University

Technology-related transformations of imaginary body boundaries: Psychopathology of the everyday excessive Internet and mobile phone use

Vadim A. Emelin[a], Elena I. Rasskazova[a,b*], Alexander Sh. Tkhostov[a]

[a] *M.V. Lomonosov Moscow State University, Moscow, Russia*

[b] *Mental Health Research Centre, Moscow, Russia*

*Corresponding author. E-mail: e.i.rasskazova@gmail.com

Background. In line with the approach of Larkin et al. (2006), we consider technological dependence in the context of the interaction between personality, environment, and culture.

Objective. The aim of this study is to discover technology-related changes in psychological needs and boundaries that could mediate the relationship between psychopathological symptoms and indicators of excessive use of info-communication technologies (ICT). The application of the Body Function Regulation Model to the use of ICT suggests that technology-related changes in the system of an individual's needs and psychological boundaries mediate the relationship between a sense of poor psychological well-being and the risk of technology dependence.

Design. The study of a normative sample (N = 275) using two technologies–mobile phones and the Internet–was performed.

Results and Discussion. We demonstrated that the relationship between the general level of psychopathological symptoms and excessive use of technology (subjective dependence and inability to refrain from use of mobile phones and the Internet) is indeed mediated by the perception of their indispensability for extension of psychological boundaries, and (for the Internet) its use in image-making.

***Keywords*:** Body function regulation model, psychological consequences of technologies, psychopathological complaints, the revised version of the Technology-Related Psychological Consequences Questionnaire, excessive use of technologies

Introduction

The terms "excessive use" and "problematic use" (Bianchi & Phillips, 2005) of technologies, as well as "technological addiction" (Griffiths, 2000, 2005), refer to an individual spending so much time using some form of info-communication tech-

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0312
http://psychologyinrussia.com

nology that he/she neglects other life spheres, or adopts a potentially dangerous use of the technology (like cyberbullying or calling when driving).

While scientists are still arguing about whether to consider technological addiction a mental disorder, and what the criteria for its definition should be (Kuss & Lopez-Fernandez, 2016), most researchers agree that this is a practical problem with rather higher incidence. The components used to define technological addiction are typically taken from the criteria used to define substance use or gambling (e.g., salience, mood modification, tolerance, withdrawal symptoms, conflict, and relapse– Griffiths, 2005), and include subjective appraisals of dependency and the negative impact on personal life spheres.

Excessive use of both the Internet and mobile phones were found to be a widespread problem, especially among adolescents and youth. For instance, in Britain, 10% of students report extensive use of mobile phones (Lopez-Fernandez et al., 2014); in China, 11.7% of student Internet users meet the criteria for Internet addiction (Li et al., 2014). Comparisons of 31 nations revealed a 6.0% (Cheng & Li, 2014) mean prevalence rate for Internet addiction, with the highest rate in the Middle East (10.9%) and the lowest rate in Northern and Western Europe (2.6%). According to a study of the Russian population, 11% of adolescents report three or more symptoms of excessive Internet use (Soldatova & Rasskazova, 2013). Internet addiction is related to high stress, sleep disturbances (Younes et al., 2016, Cannan et al., 2013), and lower subjective and objective quality of life (Cheng & Li, 2014).

There is a growing body of data showing that problematic or excessive use of info-communicational technologies is related to a higher level of psychopathological complaints and different psychological factors. Internet addiction was found to be related to anxiety, depression, and low self-esteem (Younes et al., 2016), extrinsic locus of control (Andreou & Svoli, 2013), and a wide range of psychopathological symptoms (Kuss & Lopez-Fernandez, 2016), including even psychotic-like experiences (Mittal et al., 2013). Moreover, technological addiction is accompanied by a preference for specific kinds of perception of one's self and others, and coping strategies. Adolescents with problematic use of mobile phones tend to perceive themselves as experts in this technology, and see their peers as using their phones extensively as well (Lopez-Fernandez et al., 2014).

Experiential avoidance as a self-regulatory strategy, which involves trying to control or escape from negative stimuli such as thoughts, feelings, or sensations that generate strong distress (Garcia-Oliva & Piqueras, 2016), was found to be an important factor involved in addictive use of the Internet, mobile phones, and video games in adolescents The relationship between Internet addiction and psychosocial maladjustment was found to be mediated by coping inflexibility and coping avoidance (Cheng et al., 2015). A sense of lack of perceived control, especially if accompanied by a perception of a low level of familial emotional support, predicts the ability of the Internet to negatively affect a person's life (Pace et al., 2014).

However, due to the cross-sectional design of most studies, it is unclear whether psychological changes are the reason for, or an outcome of, the extensive use of technology. Although the relationship between psychopathology and technological addiction was traditionally interpreted as an indicator of vulnerability, some researchers insisted that understanding of technological addictions demands consideration of the whole system of "personality–environment–culture" (Larkin et al., 2006) and the context of excessive use (Griffiths, 2010). According to an Interaction of Person-Affect-Cognition-Execution (I-PACE) model of specific Internet-use disorders (Brand et al., 2016), while personality factors could predispose an individual to different Internet-use disorders, cognitive and affective processes further mediate addiction formation. Thus, both reasons for and effects of addiction depend not only on the individual's personality and symptoms; rather, while the personality creates specific conditions for dependence, the dependence's development is a result of different factors. In line with this idea, there is data showing that psychopathological symptoms are different in those Internet addicts who have social dysfunction, compared with those who have not: the former have higher levels of interpersonal sensitivity, hostility, and paranoia; lower levels of social responsibility, anxiety, self-control, and family social support; and they were more likely to employ negative coping strategies (Chen et al., 2015).

The purpose of this study is to uncover technology-related psychological changes that could mediate relationship between psychopathological symptoms and indicators of excessive use of info-communication technologies. In line with the approach of Larkin et al. (2006), we consider technological dependence in the context of the interaction between personality, environment, and culture, applying the Body Function Regulation Model (Tkhostov, 2002) to the use of ICT.

Perception of information technologies as indispensable and invisible: The psychological model of body function regulation

The Body Function Regulation Model(Tkhostov, 2002) was suggested as a framework for explaining functional and behavioral disorders from a socio-cultural perspective. It suggests that the perception of objects as "mine" and part of "me" is based on a feeling of their controllability. Loss of controllability (due to too high social demands for functioning, damage, or external influence) leads to the externalization of functions that are felt to be "not mine any more." Attempts to consciously regulate them are typically unhelpful and might provoke the perpetuation of the problem. The application of this approach to ICT allows us to consider technologies not only as human extensions (MacLuhan, 1964) opening up new opportunities (Reingold, 2002), but also as psychological transformations (Emelin et al., 2012). As a result, technologies become indispensable to the person, and invisible, creating an illusion of controllability and the subjective necessity to use them, even if it is not obligatory.

Indeed there is data (Walsh & White, 2007) demonstrating that technology use is closely related to subjective appraisals of its controllability and the development of a technology-related identity (self-identity, similarity to a prototype, the emotional appraisal of the prototype). According to L. Tian et al. (2009), any technology which makes access to information quick and easy, reduces the feeling of uncertainty and increases the feeling of safety. In analyzing psychological changes in users of mobile phones, L. Srivastava (2005) suggests that they are perceived as "always necessary," making them subjectively indispensable, and their loss personally traumatic.

The application of the Body Function Regulation Model to technologies (Emelin et al., 2012) identifies different changes in psychological boundaries: the feeling of controllability (boundary extension); vulnerability to others who can always achieve (boundary violation); and preference for the technologies over activities that are not moderated by technologies. In this study we used two indicators of excessive use of two technologies (mobile phones and the Internet): subjective dependence, and the inability to refrain from using them.

The study puts forward several hypotheses:

1. Psychopathological complaints are related to indicators of excessive use of mobile phones and the Internet in the normative sample.
2. This relationship is mediated by an extension of psychological boundaries and perceived changes in needs due to the technologies.
3. The mediation effects do not depend on the frequency of technology use.

Method

In accordance with the study's objectives, we sought a broad adult sample (not only youth). To obtain this sample, second-year students of the psychology faculty were asked to advertise this research project in their communities. Inclusion criteria were: 1) an age range of 17–70 years old; 2) access to and experience of use of the Internet and mobile phones (even if they don't use them now); and 3) the absence of diagnosed mental illness. The volunteer subjects then filled out questionnaires. While it is typical in Russia for women and younger people to be more open to participating in psychological studies, recruitment continued until at least 100 males were included. Two hundred and seventy-four people living in Moscow and the Moscow region (100 males, 174 females, mean age 25.8 ± 11.8) participated in the study. Of these, 95 (34.7%) had higher education; others had completed basic education, and were either students (72 participants, 26.3%) or worked (107 participants, 39.1%). Seventy-five (27.3%) were married or lived with a partner, 187 (68.3%) were not married, and 12 (4.4%) were divorced. Sixty-three (23.0%) had children.

The study was of the cross-sectional design. Participants supplied the following information:

1. **The revised version of Technology-Related Psychological Consequences Questionnaire**.[1] This is a screening instrument (Emelin et al., 2014) based on the body function regulation approach (Tkhostov, 2002), and applies to both technologies (the Internet and communication by mobile phones). There are nine scales within each technology, which are divided into three blocks: indicators of excessive use, psychological boundary transformation, and needs transformation. Each scale is tested by three items appraised on the 4-item Likert scale. The block of indicators of excessive use includes two scales: the inability to refrain from use (e.g., "I can't imagine my life without mobile phone") and subjective dependence (e.g.,

[1] The revised version of Technology-Related Psychological Consequences Questionnaire and further details on its structure and items are available from the corresponding author.

"I spend more time on the Internet than I would like to"). The block of transformation of boundary extension includes four scales: 1) boundary extension (e.g., "If the person with whom I used to talk in the Internet isn't online for a long time, I worry"); 2) boundary violation (e.g., "I'm concerned that my personal information may be available to anyone in the Internet"); 3) easiness-related preference for the technology (e.g., " The Internet can substitute for lots of hobbies and real-life activities"); and 4) opportunity-related preference for technologies (e.g., "It's important that the mobile phone makes it easy for me to distract myself from unpleasant discussion or events").

Technology-related needs each have three scales: 1) functionality ("I like the fact that with the Internet I can at any time send a message to any people I need to contact wherever I am or they are"); 2) convenience ("All I need in a computer is for it to be reliable and easy to use"); and 3) image-making ("I prefer to buy an expensive but stylish mobile phone"). In this study we didn't use the boundary violation scale, due to ambiguity in its interpretation (Emelin et al., 2014): it reflects both the degree of violation and sensitivity to the impact of the technologies.

2. **Frequency of use** of the Internet and mobile phone was assessed by the response to one statement on a 4-point Likert scale: "Typically I use … (the Internet or mobile phone) never or almost never / rarely / sometimes / often".
3. **Symptom Checklist-90-Revised** (Derogatis & Salvitz, 2000) includes a list of 90 psychopathological symptoms measuring Somatization, Obsessive-Compulsiveness, Interpersonal Sensitivity, Depression, Anxiety, Hostility, Phobic Anxiety, Paranoid Ideation, and Psychoticism. There are three secondary global indexes: the Global Severity Index (measuring overall psychological distress); the Positive Symptom Distress Index (measuring the intensity of symptoms); and the Positive Symptom Total (measuring the number of self-reported symptoms). In line with data showing that there are general psychopathological complaints which increase with different psychiatric conditions (e.g., demoralization, Tellegen et al., 2008), the scores across the nine scales were highly consistent (Cronbach's alpha .95). Thus we constructed a mean score of psychopathological complaints across all the scales of SCL-90R, thus measuring a general level of vulnerability to psychopathology.

Taking into account the imbalance in gender and age in the sample, we tested for their effects on all the dependent variables. Gender was unrelated to both the inability to refrain from, and subjective dependence on, the use of mobile phones and the Internet. Age was unrelated to the inability to refrain from mobile phone use ($r = -.07$), but as expected, negatively correlated to subjective dependence, as well as to the inability to refrain from Internet use ($r = -.29$ to $-.18$). Nevertheless, to be sure that the results were not distorted by the effects of age and gender, all the analyses were repeated with the inclusion of both factors as covariates. As shown below, the general pattern of results remained the same. In the discussion section, we further address the possible limitations due to this imbalance.

Results

Indicators of excessive use of technologies and psychopathological complaints

There were medium positive correlations between subjective dependence and the inability to refrain from use of the technologies ($r = .25$, $p < .01$ for mobile phones and $r = .41$ for the Internet) and across technologies ($r = .30–.36$, $p < .01$)[1]. Both indicators of excessive technologies use were related to more frequent use ($r = .24–.39$ for mobile phones and $r = .27$ for Internet)[2].

Table 1. Correlations between psychopathological complaints (SCL-90-R), subjective dependence and the inability to refrain from mobile phone and Internet use. (Partial correlations after adjusting for age and gender are in the parentheses).

SCL-90-R Scales and Indexes	Inability to refrain from mobile phone use	Subjective dependence on mobile phone use	Inability to refrain from Internet use	Subjective dependence on the mobile phone and Internet
Somatization	.10 (.13*)	.23** (.24**)	.02(.01)	.19** (.18**)
Obsessive-Compulsiveness	.19** (.19**)	.29** (.26**)	.19** (.14*)	.38** (.32**)
Interpersonal Sensitivity	.19** (.18**)	.22** (.18**)	.20** (.16**)	.35** (.31**)
Depression	.16** (.15*)	.21** (.18**)	.15* (.10)	.36** (.30**)
Anxiety	.20** (.20**)	.23** (.21**)	.13* (.09)	.31** (.27**)
Hostility	.14* (.14*)	.21** (.18**)	.12 (.07)	.24** (.17**)
Phobic Anxiety	.09 (.06)	.20** (.17**)	.16** (.12)	.25** (.21**)
Paranoid Ideation	.21** (.22**)	.25** (.22**)	.16** (.13*)	.30** (.26**)
Psychoticism	.19** (.20**)	.28** (.25**)	.15* (.10)	.32** (.27**)
Global Severity Index	.19** (.20**)	.28** (.25**)	.16** (.11)	.36** (.31**)
Positive Symptom Total	.15* (.16*)	.29** (.27**)	.15* (.11)	.36** (.33**)
Positive Symptom Distress Index	.20** (.20**)	.20** (.16**)	.15* (.11)	.28** (.21**)
Mean level of psychopathological complains	.20** (.20**)	.28** (.25**)	.17** (.12*)	.36** (.31**)

* — $p < .05$, ** — $p < .01$.

[1] Partial correlations after adjusting for age and gender are $r = .24$, $r = .36$ and $r = .30–.33$, consequently.

[2] Partial correlations after adjusting for age and gender are $r = .21–.37$ and $r = .18–.21$, consequently.

In line with theories of technological addictions, both subjective dependence and the inability to refrain from use were weakly positively correlated with different psychopathological complaints (Table 1): especially complaints of obsessive-compulsive symptoms, anxiety, paranoid ideation, and psychoticism. With the exclusion of the inability to refrain from Internet use, the weakest correlations disappear after adjusting for age and gender; for the other three dependent variables, the correlational pattern remains relatively stable even when age and gender are taken into account.

Indicators of excessive use of technologies and psychological boundary and needs transformation

Boundary extension and technology-based image-making were the strongest predictors of excessive use for both mobile phones and the Internet (Table 2). Easiness-related preference for technology correlated mainly with Internet use while opportunity-related preference correlated with mobile phone use. The general pattern of the results remained after adjusting for age and gender.

Table 2. Correlations between indicators of psychological boundary and needs transformation, subjective dependence and the inability to refrain from mobile phone and Internet use. (Partial correlations after adjusting for age and gender are in the parentheses).

Indicators of psychological boundary and needs transformation	Inability to refrain from mobile phone use	Subjective dependence on mobile phones	Inability to refrain from Internet use	Subjective dependence on mobile phones and Internet
Boundaries extension	.43** (.42**)	.21** (.21**)	.47** (.45**)	.47** (.45**)
Easiness-related preference of technology	.10 (.07)	.15* (.12*)	.32** (.31**)	.30** (.30**)
Opportunity-related preference of technology	.40** (.42**)	.47** (.46**)	.20** (.20**)	.11 (.11)
Functionality	.34** (.35**)	-.02 (.00)	.18** (.15*)	.19** (.14*)
Convenience	.23** (.22**)	.02 (-.04)	.21** (.20**)	.16** (.14*)
Image making	.46** (.48**)	.37** (.35**)	.41** (.38**)	.46** (.42**)

* — $p<.05$, ** — $p<.01$.

Psychological boundaries and needs transformation as mediators of the relationship between psychopathological complaints and indicators of excessive use of technologies

To test the hypothesis that psychological transformations mediate the relationship between psychopathological complaints and indicators of excessive use of technologies, we conducted a number of mediation analyses separately for mo-

Table 3. Psychological boundary and needs transformation as mediators of the relationship between psychopathological complaints and indicators of excessive use of mobile phone and Internet. (Effects after control for age and gender are in parentheses).

Mediator	β Complaints — Mediator	β Mediator – DV	Direct effect β Complaints – DV [95% CI]	Indirect effect β Complaints – DV [95% boot-strapped 10,000 CI]
Dependent variable: Inability to refrain — Mobile phones				
Boundaries extension	.51** (.51**)	.41** (.38**)	.14 [–.06 – .34] (.14 [–.06 – .34])	.21 [.12 – .32] (.19 [.11– .30])
Opportunity-related preference of technology	.44** (.46**)	.37** (.38**)	.19 [–.01 – .39] (.16 [–.04 – .36])	.16 [.08 – .28] (.18 [.09 – .29])
Dependent variable: Subjective dependence - Mobile phones				
Boundaries extension	.51** (.51**)	.14* (.15*)	.43** [.21 – .64] (.36** [.15 – .58])	.07 [.01 – .15] (.07 [.02 – .16])
Opportunity-related preference of technology	.44** (.46**)	.43** (.41**)	.31** [.12 – .50] (.25* [.05 – .44])	.19 [.11 – .30] (.19 [.11 – .30])
Dependent variable: Inability to refrain - Internet				
Boundaries extension	.69** (.69**)	.48** (.46**)	–.03 [–.23 – .17] (–.10 [–.31 – .11])	.33 [.21 – .48] (.32 [.20 – .47])
Easiness-related preference of technology	.37** (.39**)	.30** (.30**)	.19 [–.01 – .40] (.11 [–.10 – .31])	.11 [.05 – .19] (.11 [.05 – .20])
Image making	.42** (.42**)	.39** (.35**)	.14 [–.06 – .34] (.07 [–.13 – .28])	.16 [.08 – .28] (.15 [.07 – .25])
Dependent variable: Subjective dependence - Internet				
Boundaries extension	.69** (.69**)	.39** (.37**)	.37** [.17 – .57] (.29** [.08 – .50])	.27 [.17 – .40] (.25 [.15 – .38])
Easiness-related preference of technology	.37** (.39**)	.24** (.24**)	.55** [.36 – .75] (.45** [.25 – .66])	.09 [.03 – .17] (.09 [.04 – .18])
Image making	.42** (.42**)	.40** (.34**)	.48** [.29 – .66] (.40** [.21 – .60])	.17 [.09 – .27] (.14 [.08 – .24])

* — $p < .05$, ** — $p < .01$.

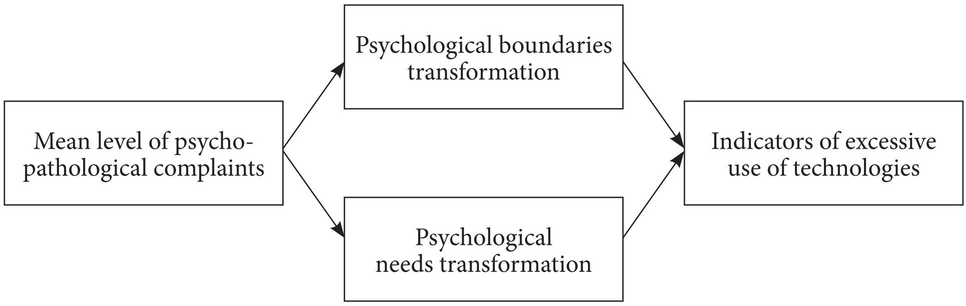


Figure 1. Schema of mediation analysis

bile phones and the Internet, and for both indicators (Chaplin, 2007). The dependent variables were the subjective dependence on technology, and the inability to refrain from its use, and the independent variable was the mean level of psychopathological complaints (Fig. 1). Three scales of the psychological boundary transformation block, and three scales of psychological needs transformation block, were tested as possible mediators. Preacher and Hayes' macro for SPSS was used to measure indirect bootstrapped effects[1]. Unstandardized indirect effects were computed for 10,000 bootstrapped samples, and the 95% confidence intervals were obtained (Table 3).

For both subjective dependence and the inability to refrain from mobile phone use, the effect of the psychopathological complaints was mediated by boundaries extension and the opportunity-related preference for technology. No direct effect of psychopathological complaints on the inability to refrain from mobile phone use was found, allowing us to hypothesize that vulnerability to mobile phone-related psychological transformations in people with higher psychopathological complaints is a key factor in their inability to refrain from mobile phone use.

Furthermore, both the effects of psychopathological complaints on subjective dependence, and the inability to refrain from Internet use are mediated by boundary extension, easiness-related preference for the Internet, and image-making online. As in the case of mobile phones, we found no significant direct effects of psychopathological complaints on the inability to refrain from Internet use.

The same mediation patterns remained after we added the frequency of use variable in the first step of regression analysis. Moreover, as can be clearly seen in the table, adjusting for age and gender leads to minor changes in the results.

Discussion

According to our data, higher psychopathological complaints are indeed related to higher subjective dependence, and the inability to refrain from the use of the technologies (both mobile phones and the Internet). However, these effects are mediated by technology-related changes in psychological boundaries and needs. Specifically, people with higher psychopathological complaints tend to feel that they can control and achieve more than others by using mobile phones and the Internet,

[1] URL: http://processmacro.org/download.html

and this feeling, if developed, contributes to the risk of subjective dependence and the inability to refrain from use of the technologies. Moreover, these people more frequently consider mobile phones as indispensable due to the opportunities they open up, while the Internet is seen as indispensable due to its easiness. These feelings are consequently related to indicators of excessive use of mobile phones and the Internet.

In general, these results are in accordance with vulnerability-stress-coping model that has been suggested for mental illnesses (especially schizophrenia, see Zubin & Spring 1977), in that external stimuli (e.g., technology use) can amplify initial personal vulnerabilities. Moreover, the data indicate that psychological complaints should be considered a risk factor creating vulnerability to technology-related changes, but not to their excessive use per se (Larkin et al., 2006). An interesting question for future research is why psychopathological symptoms are related to a higher vulnerability to psychological transformations when using technologies.

Technology-related transformation in needs seem to be unrelated to the "psychopathological complaints–indicators of excessive use" relationship. The only exception refers to image-making: people with higher psychopathological complaints more often believe that Internet improves their image, and this belief predicts both subjective dependence and the inability to refrain from its use. L. Srivastava (2005) suggests the construct of "personalization" of mobile phones that make them subjectively indispensable, and their loss personally traumatic. She describes (e.g., Srivastava, 2005) changes in psychological needs when technologies obtain some additional meanings for the person (e.g., "to have an expensive mobile phone" means "to look decent").

Compared to mobile phone use, which uses "real-life" communication to make a person's image, the image-making function of the Internet occurs "through" the Internet. We can hypothesize that the preference for online image-making is sometimes due to problematic image-making offline, or due to the choice of an online image that is socially undesirable and could lead to poorer adjustment. Supporting this hypothesis is the fact that youth preference for online self-presentation is associated with a less stable sense of self (Fullwood et al., 2016), although longitudinally it could lead to better self-esteem (Yang & Brown, 2016). It could be that (both in youth and adults) by becoming the main sphere where person is able to create and share his preferred meaning with others, the Internet becomes a highly unique place, thus increasing the person's vulnerability to Internet addiction.

In accordance with the psychological model of body function regulation in this study, we aimed to reveal relationships that are common for males and females and for different ages (e.g., not only youth). Although adjusting for age and gender didn't change the whole pattern of results, it should be noted that our sampling strategy (resulting in prevalence of females and younger people) could potentially lead to poor representation of older people (55+). This group demands further study on special samples. We still believe that the results could be extended to people of both genders and to both youth and adults in their 30s to 50s, because these groups are represented in a volume that justifies their statistical control. Certainly —as in most psychological research —it is still possible that people who agreed to participate in the study differ from those who refused.

In general, our data support the idea that experiences of controllability and indispensability could make an important contribution to explaining technology use and technology-related adaptation.

Acknowledgements

This study was supported by the Russian Foundation for Humanitarian Sciences, project 14-06-00730.

References

Andreou, E. & Svoli, H. (2013). The association between Internet user characteristics and dimensions of Internet addiction among Greek adolescents. *International Journal of Mental Health and Addiction*, *11*, 139–148. doi: 10.1007/s11469-012-9404-3

Bianchi, A. & Phillips, J.C. (2005). Psychological predictors of problem mobile phone use. *Cyberpsychology and Behavior*, *8*, 39–51. doi: 10.1089/cpb.2005.8.39

Brand, M., Young, K.S., Laier, C., Woefling, K., & Potenza, M.N. (2016). Integrating psychological and neurobiological considerations regarding the development and maintenance of specific Internet-use disorders: An Interaction of Person-Affect-Cognition-Execution (I-PACE) model. *Neuroscience & Biobehavioral Reviews, 71*, 252–266. doi: 10.1016/j.neubiorev.2016.08.033

Canan, F., Yildirim, O., Sinani, G., Ozturk, O., Ustinel, T.Y., & Ataoglu, A. (2013). Internet addiction and sleep disturbance symptoms among Turkish high school students. *Sleep and Biological Rhythms, 11*, 210–216. doi: 10.1111/sbr.12022

Chaplin, W.F. (2007). Moderator and mediator models in personality research. In R.W. Robins, R.C. Fraley & R.E. Krueger (Eds.), *Handbook of research methods in personality psychology* (pp. 602–632). New York: The Guilford Press.

Chen, Q., Quan, X., Lu, H., & Fei, P. (2015). Comparison of the personality and other psychological factors of students with internet addiction who do and do not have associated social dysfunction. *Shanghai Archives of Psychiatry*, *27*(1), 36–41. doi: 10.11919/j.issn.1002-0829.214129

Cheng, S. & Li, A.Y. (2014). Internet addiction prevalence and quality of (real) life: A meta-analysis of 31 nations across seven world regions. *Cyberpsychology, Behavior and Social Networking*, *17*(12), 755–760. doi: 10.1089/cyber.2014.0317

Cheng, C., Sun, P., & Mak, K. (2015). Internet addiction and psychosocial maladjustment: Avoidant coping and coping inflexibility as psychological mechanisms. *Cyberpsychology, Behavior and Social Networking*, *18*(9), 539–546. doi: 10.1089/cyber.2015.0121

Derogatis, L.R. & Savitz, K.L. (2000). The SCL-90-R and the Brief Symptom Inventory (BSI) in primary care. In M. E. Maruish (Ed.), *Handbook of psychological assessment in primary care settings*. (Vol. 236, pp. 297–334). Mahwah, NJ: Lawrence Erlbaum Associates.

Emelin, V.A., Tkhostov, A.Sh., & Rasskazova, E.I. (2014). Psychological adaptation in the info-communication society: The revised version of the Technology-Related Psychological Consequences Questionnaire. *Psychology in Russia: State of the Art,* 7(2), 105–120. doi: 10.11621/pir.2014.0210

Emelin, V.A., Rasskazova, E.I., & Tkhostov, A.Sh. (2012). Psikhologicheskie posledstviya razvitiya informatsionnykh tekhnologii [Psychological consequences of informational technologies development]. *Natsionalnyi psikhologicheskii zhurnal [National Psychological Journal], 1,* 81–87.

Fullwood, C., James, B.M., & Chen-Wilson, C.-H. (2016). Self-concept clarity and online self-presentation in adolescents. *CyberPsychology, Behavior & Social Networking*, *19*(12), 716–720. doi: 10.1089/cyber.2015.0623

García-Oliva, C. & Piqueras, J. (2016). Experiential avoidance and technological addictions in adolescents. *Journal of Behavioral Addictions*, *5*(2), 293–303. doi: 10.1556/2006.5.2016.041

Griffiths, M. (2010). The role of context in online gaming excess and addiction: some case study evidence. *International Journal of Mental Health and Addiction*, *8*, 119–125. doi: 10.1007/s11469-009-9229-x

Griffiths, M. (2005). A 'components' model of addiction within biopsychosocial framework. *Journal of Substance Use*, 10(4), 191–197. 10.1080/14659890500114359

Griffiths, M. (2000). Internet addiction–time to be taken seriously. *Addiction Research*, 8(5), 413–418. 10.3109/16066350009005587

Kuss, D.J. & Lopez-Fernandez, O. (2016). Internet addiction and problematic Internet use: A systematic review of clinical research. *World Journal of Psychiatry, 6*(1), 143–176. doi: 10.5498/wjp.v6.i1.143

Larkin, M., Wood, R.T., & Griffiths, M.D. (2006). Toward addiction as relationship. *Addiction Research and Theory*, *14*(3), 207–215. doi: 10.1080/16066350500151747

Li, Y., Zhang, X., Lu, F., Zhang, Q., & Wang, Y. (2014). Internet addiction among elementary and middle school students in China: A nationally representative sample study. *Cyberpsychology, Behavior and Social Networking*, *17*(2), 111–116. doi: 10.1089/cyber.2012.0482

Lopez-Fernandez, O., Honrubia-Serrano, L., Freixa-Blanxart, M., & Gibson, W. (2014). Prevalence of problematic mobile phone use in British adolescents. *Cyberpsychology, Behavior and Social Networking*, *17*(2), 91–98. doi: 10.1089/cyber.2012.0260

McLuhan, M. (1964). *Understanding media: The extensions of man.* 1st ed. New York: McGraw Hill.

Mittal, V.A., Dean, D.J., & Pelletier, A. (2013). Internet addiction, reality substitution and longitudinal changes in psychotic-like experiences in young adults. *Early Intervention in Psychiatry, 7*, 261–269. doi: 10.1111/j.1751-7893.2012.00390.x

Pace, U., Zappulla, C., Guzzo, G., Di Maggio, R., Laudani, C., & Cacioppo, M. (2014). Internet addiction, temperament, and the moderator role of family emotional involvement. *International Journal of Mental Health and Addiction, 12*, 52–63. doi: 10.1007/s11469-013-9468-8

Rheingold, H. (2002). *Smart mobs: The next social revolution*. New York: Basic Books.

Soldatova, G.U. & Rasskazova, E.I. (2013). Excessive Internet use: Factors and indicators. *Psikhologicheskii zhurnal [Psychological Journal]*, *34*(4), 105–114.

Srivastava, L. (2005). Mobile phones and the evolution of social behavior. *Behavior and Information Technology*, *24*(2), 111–129. doi: 10.1080/01449290512331321910

Tian, L., Shi, J., & Yang, Z. (2009). Why does half the world's population have a mobile phone? An examination of consumers' attitudes toward mobile phones. *Cyberpsychology and Behavior*, *12*(5), 513–516. doi: 10.1089/cpb.2008.0335

Tellegen, A., Ben-Porath, Y.S., McNulty, J.L., Arbisi, P.A., Graham, J.R., & Kaemmer, B. (2008). *The MMPI-2 Restructured Clinical scales: Development, validation, and interpretation*. Minneapolis: University of Minnesota Press.

Tkhostov, A.Sh. (2002). *Psikhologiya telesnosti* [Psychology of body function regulation]. Moscow: Smysl.

Walsh, S.P. & White, K.M. (2007). Me, my mobile and I: The role of self- and prototypical identity influences in the prediction of mobile phone behavior. *Journal of Applied Social Psychology*, *37*(10), 2405–2434. doi: 10.1111/j.1559-1816.2007.00264.x

Yang, C. & Brown, B. (2016). Online self-presentation on Facebook and self development during the college transition. *Journal of Youth & Adolescence*, *45*(2), 402–416. doi: 10.1007/s10964-015-0385-y

Younes, F., Halawi, G., Jabbour, H., Osta, N.E., Karam, L., Hajj, A., & Khabbaz, L.R. (2016). Insomnia, anxiety, depression, stress and self-esteem in university students: A cross-sectional designed study. *PLOS One, 11*(9), 13. doi: 10.1371/journal.pone.0161126

Zubin, J. & Spring, B. (1977). Vulnerability — a new view of schizophrenia. *Journal of Abnormal Psychology*, *86*, 103–126. doi: 10.1037/0021-843X.86.2.103

Original manuscript received February 03, 2017
Revised manuscript accepted May 18, 2017
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society | Lomonosov Moscow State University

ADVANCES IN COGNITIVE NEUROSCIENCE

Not all reading is alike: Task modulation of magnetic evoked response to visual word

Anna A. Pavlova[a, c*], Anna V. Butorina[a], Anastasia Y. Nikolaeva[a], Andrey O. Prokofyev[a], Maxim A. Ulanov[a, b], Tatiana A. Stroganova[a]

[a] *Moscow State University of Psychology and Education, Moscow, Russia*

[b] *National Research University Higher School of Economics, Moscow, Russia*

[c] *Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia*

* Corresponding author. E-mail: anne.al.pavlova@gmail.com

Background. Previous studies have shown that brain response to a written word depends on the task: whether the word is a target in a version of lexical decision task or should be read silently. Although this effect has been interpreted as an evidence for an interaction between word recognition processes and task demands, it also may be caused by greater attention allocation to the target word.

Objective. We aimed to examine the task effect on brain response evoked by non-target written words.

Design. Using MEG and magnetic source imaging, we compared spatial-temporal pattern of brain response elicited by a noun cue when it was read silently either without additional task (SR) or with a requirement to produce an associated verb (VG).

Results. The task demands penetrated into early (200-300 ms) and late (500-800 ms) stages of a word processing by enhancing brain response under VG versus SR condition. The cortical sources of the early response were localized to bilateral inferior occipitotemporal and anterior temporal cortex suggesting that more demanding VG task required elaborated lexical-semantic analysis. The late effect was observed in the associative auditory areas in middle and superior temporal gyri and in motor representation of articulators. Our results suggest that a remote goal plays a pivotal role in enhanced recruitment of cortical structures underlying orthographic, semantic and sensorimotor dimensions of written word perception from the early processing stages. Surprisingly, we found that to fulfil a more challenging goal the brain progressively engaged resources of the right hemisphere throughout all stages of silent reading.

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0313
http://psychologyinrussia.com

Conclusion. Our study demonstrates that a deeper processing of linguistic input amplifies activation of brain areas involved in integration of speech perception and production. This is consistent with theories that emphasize the role of sensorimotor integration in speech understanding.

Keywords: visual word recognition, top-down modulations, sensorimotor transformation, speech lateralization, magnetoencephalography (MEG)

Introduction

Visual word recognition lies at the heart of written language capacity. However, the usage of recognized information can vary greatly: a reader can pronounce the written word out loud, respond to it with another word or an action, or even scan through it without conscious access to its meaning. For a long period of time the dominant assumption in psycholinguistics has been that word recognition is unaffected by intention of the reader, in a sense that it is triggered automatically and obligatory, regardless of the task demands (see, e. g. Neely, & Kahan, 2001; Posner, & Snyder, 1975). Evidence to support this view comes from the Stroop effect (Stroop, 1935): it takes more time to name the color in which a word is written when the word and the color name conflict (e.g., the word red displayed in green font) compared to when the word is neutral with respect to color (e.g., book displayed in green font). Apparently, meaning of the words is activated despite being task-irrelevant or even disruptive for the task performance (see, e.g., MacLeod, 1991, 2005, Velichkovsky, 2006, for reviews).

The multiple priming studies also showed that a visually presented word elicits automatic access to words and their meaning (e.g., Forster, & Davis, 1984; Marcel, 1983). The priming refers to the consistent finding that processing of a word is facilitated if it is preceded by semantically related prime word (e.g., cat–dog) relative to when it is preceded by semantically unrelated word (e.g., ball–chair). If the prime is presented briefly (e.g. 30 ms) and immediately replaced by a mask (e.g. a pattern of symbols at the same spatial location as the prime), participants are usually unable to report having seen the prime but, nonetheless, respond faster and more accurate if it is semantically related to the target (see e.g., Carr, & Dagenbach, 1990; Marcel, 1983; Neely, 1991; Neely, & Kahan, 2001). Accepting that the mask procedure prevents the consolidation of long-lasting episodic memories and, therefore, any top-down guidance of word recognition, the masked priming was considered as a key evidence in favor of the theoretical view that language-processing system is an insulated cognitive module impervious to top-down control modifications (Fodor, 1983).

However, over the last decades new data challenged the claim of strict automaticity of word recognition. Several behavioral studies showed that the strength of Stroop effect depends on the context (e.g., Balota, & Yap, 2006; Besner, 2001; Besner, Stolz, & Boutilier, 1997). If only a single letter in a Stroop word is colored (rather than all the letters), or a single letter within a word is spatially pre-cued (rather than whole word), or a ratio of congruent–incongruent trials is low (20:80), the Stroop effect is significantly reduced (e.g. Besner, 2001). Moreover, there is evidence that the effects of masked primes and primes presented with very short prime-target SOA also can be modulated by such top-down factor as expectancy: decreasing the proportion of related primes leads to smaller priming effect (Balota, Black, & Cheney, 1992; Bodner, & Masson, 2004).

Non-automatic nature of word perception was also evidenced by the findings demonstrating that the goal of word-related task changed which characteristics of the perceived words were important for the performance. While word naming speed depends on word's phonological onset variables (voicing, location, and manner of articulation of word first phonemes), the speed of lexical or semantic decision regarding the same words has been mostly influenced by their lexical frequency or imageability (Balota, Burgess, Cortese, & Adams, 2002; Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Balota & Chumbley, 1984; Ferrand et al., 2011; Kawamoto, Kello, Higareda, & Vu, 1999; Schilling, Rayner, & Chumbley, 1998; Yap, Pexman, Wellsby, Hargreaves, & Huff, 2012).

Yet, the behavioral evidence alone cannot distinguish between two alternative ways the task may modulate word processing. One option is that the pursued goal affects word recognition, i.e. the way relevant information is retrieved from a written word. Alternatively, the goal may affect the following stage of decision making determining how to use the retrieved information to achieve the intended result. The information of exactly which stages of word recognition are affected could be obtained from electrophysiological studies which allow to detect exact timing when word recognition process is penetrated by task demand.

Electrophysiologically, the majority of research addressing the task effects on the brain response to written word was concentrated on the N400 component of event-related potentials (ERP) — a negative deflection that peaks at approximately 400 ms after stimulus-onset and is thought to reflect the processing of word semantics. These studies demonstrated that stronger emphasis placed on semantic attributes of words by the task demands enhanced N400 suggesting that a goal penetrates into word semantic processing (Bentin, Kutas, & Hillyard, 1993, 1995).

The recent neurophysiological studies have demonstrated that task demands can modulate word recognition at much earlier latencies than N400 time window. Ruz and Nobre (2008) in a cuing paradigm showed that a negative ERP deflection with peak at 200 msec post-stimulus was larger when the attention was oriented toward orthographic rather than phonological attributes of words. Strijkers with colleagues (2011) reported that the brain electrical response in reading aloud versus semantic categorization task starts to dissociate around 170 msec after a word onset. Similarly, Mahé, Zesiger and Laganaro (2015) observed early differences between lexical decision and reading aloud at ERP waveform from ~180 ms to the end of the analyzed interval (i.e., 500 ms). Chen and colleagues (2013) using MEG found that different task sets (lexical decision, semantic decision and silent reading) affect the word processing already at first 200 msec after word presentation. Stronger activation was observed for lexical decision in areas involved in orthographic and semantic processes (such as left inferior temporal cortex and bilateral anterior temporal lobe) Thus, early modulations of the word-related ERP are especially pronounced, if the task demands direct attention towards semantic or lexical features of the perceiving word. The early time window, where these effects occurred are theoretically consistent with the results of other studies demonstrating the early onset of "rudimentary" semantic analysis of the presented word (e.g. Kissler, Herbert, Winkler, & Junghofer, 2009).

Summing up, the discussed results provide substantial support for the view that even the earliest stages of word recognition are susceptible to top-down modulations when the presented word is a target for the task in hand. In the current study we aimed to examine whether the word processing is affected by the task when the presented word is just a cue which triggers a subsequent memory search for the target word. We compared brain response elicited by silent reading of a noun cue either without additional task (SR) or with a requirement to further produce an action verb associated with the cue (verb generation, VG). In contrast to the paradigms used in the previous studies, here the presented word is not an immediate processing target. Therefore, our experimental design enables us to reduce unspecific and very powerful effects of selective attention to a target word, which may well explain all the previously obtained results. Using MEG and distributed source estimation procedure, we aimed to characterize how intention to produce a related word changes the recognition process of the input word and triggers computation of articulatory output.

Method

Participants

Thirty-five volunteers (age range 20–48, mean age 26, 16 females) underwent MEG recording. All participants were native Russian-speakers, right-handed, had normal or corrected-to-normal vision and reported no neurological diseases or dyslexia. Two subjects were subsequently excluded from the analysis due to MEG acquisition error and another one due to insufficient quantity of correct responses in verb generation task. The final sample comprised 32 subjects. The study was approved by the ethical committee of the Moscow State University of Psychology and Education.

Materials

One hundred thirty Russian nouns were selected as stimuli for silent reading and verb generation tasks based on the criteria that the words were concrete and contained between 4 and 10 letters. The average word length was 5.7 letters. The word form frequency was obtained from Lyashevskaya and Sharov's frequency dictionary (2009) and the average was 49.9 ipm.

Design and procedure

The participants were visually presented with the noun cues split into 14 blocks of 8 nouns each and 2 blocks containing 9 nouns. The cues within a block were randomized and presented in white font on a black background and presented on a screen placed at 1.5 m in front of the participant. The size of the stimuli did not exceed 5° of visual angle. The experiment was implemented in the Presentation software (Neurobehavioral Systems, Inc., Albany, California, USA).

Each noun was presented within two different experimental sessions. Within SR session the participant's task was to read words inwardly. During reading task the stimulus was presented for 1000 ms, and the white fixation cross preceding the cue for 300 ms with a jitter between 0-200 ms. Within VG session a participant was required to produce the verb associated with a presented noun by answering the question what this noun does. For verb generation task each noun remained on the screen for 3500 ms and was preceded by white fixation cross presented for 300–500 ms.

In VG session participants' vocal responses were tape recorded and checked for response's errors. The trials with semantically unrelated responses, incomprehensible verbalizations, imprecise vocalization onsets, and with pre-stimulus intervals overlapped with the vocal response to the previous stimulus were excluded from further analyses. As the verb responses were meant to be inflected for person and number and could be put into the reflexive form, we considered semantically correct but erroneously inflected verbs (e.g. "kvartyra - ubirayet/an apartment cleans" instead "kvartyra ubirayetsya/an apartment is cleaned") as errors and also removed them from the subsequent analysis.

MEG data acquisition

MEG data were acquired inside a magnetically shielded room (AK3b, Vacuumschmelze GmbH, Hanau, Germany) using a dc-SQUID Neuromag™ Vector View system (Elekta-Neuromag, Helsinki, Finland) comprising 204 planar gradiometers and 102 magnetometers. Data were sampled at 1000 Hz and filtered with a band-passed 0.03-333 Hz filter. The participants' head shapes were collected with a 3Space Isotrack II System (Fastrak Polhemus, Colchester, VA) by digitizing three anatomical landmark points (nasion, left and right preauricular points) and additional randomly distributed points on the scalp. During the recording, the position and orientation of the head were monitored by four Head Position Indicator (HPI) coils. The electrooculogram (EOC) was registered with two pairs of electrodes located above and below the left eye and at the outer canthi of both eyes for recording of vertical and horizontal eye movements respectively. Structural MRIs were acquired for 28 participants with a 1.5 T Philips Intera system and were used for reconstruction of the cortical surface using Freesurfer software (http://surfer.nmr.mgh.harvard.edu/). Head models for the rest five participants failed to be obtained because of MRI acquisition error.

MEG pre-processing

The raw data were subjected to the temporal signal space separation (tSSS) method (Taulu, Simola, & Kajola, 2005) implemented in MaxFilter program (Elekta Neuromag software) aimed to suppress magnetic interference coming from sources distant to the sensor array. Biological artifacts (cardiac fields, eye movements, myogenic activity) were corrected using the SSP algorithm implemented in Brainstorm software (Tadel, Baillet, Mosher, Pantazis, & Leahy, 2011). To countervail for within-block head-movement (as measured by HPI coils) movement compensation procedure was applied. For sensor-space analysis, the data were converted to standard head position ($x = 0$ mm; $y = 0$ mm; $z = 45$ mm) across all blocks.

Data were divided into epochs of 1500 ms, from 500 ms before up to 1000 ms after stimulus onset. Epochs were rejected if the peak-to-peak value over the epoch exceeds $3 \times 10–10$ T/m (gradiometers) and $12 \times 10–10$ T/m (magnetometers) channels.

MEG data analysis

The difference in the magnitude of noun-evoked response during silent reading versus verb generation task was examined using Statistical Parametric Mapping software (SPM12: Wellcome Trust Centre for Neuroimaging, London, http://www.fil.ion.ucl.ac.uk/spm). For analysis of evoked magnetic fields the planar gradiometers data were converted to a Matlab-based, SPM format and baseline corrected over –350 – –50 prestimulus interval. The epoched data were averaged within each task separately, using a SPM built-in averaging procedure (Holland, & Welsch, 1977). The averaged data from the each pair of planar gradiometers were combined by calculating the root-mean-square values. The resulting 3D files of space (32 × 32 pixels) and time (1000 ms) dimensions were converted to images of Neuroimaging Informatics Technology Initiative (NIfTI) format.

For statistical analysis the topography x time images were smoothed in space-time using a Gaussian smoothing kernel with Full Width Half Maximum of 8 mm × 8 mm × 8 ms to ensure that the images conform to the assumptions of Random Field Theory (Kilner, & Friston, 2010). Then, the smoothed images from SR and VG tasks were subjected to a paired t-test. The resulting statistical parametric maps underwent the false discovery rate (FDR) correction with cluster-level threshold of $p < 0.05$. The clusters that survived cluster-level correction were used to guide the subsequent analysis in the source space.

The cortical sources of the evoked responses were modelled by a "depth-weighted" linear L2-minimum norm estimation method (Hämäläinen, & Ilmoniemi, 1994) implemented in Brainstorm software (Tadel et al., 2011). Only those 28 participants for whom MRI scans were obtained entered the source analysis. The individual ERFs for each task were computed by averaging the trials within the condition over a 350 ms prestimulus interval and a 1000 ms post-stimulus for each of the 306 sensors. The cortical sources of the evoked responses were modelled by a "depth-weighted" linear L2-minimum norm estimation method (Hämäläinen & Ilmoniemi, 1994). The individual cortical surfaces were imported from FreeSurfer and tessellated with 15 000 nodes. The forward solution was calculated using overlapping spheres approach (Huang, Mosher, & Leahy, 1999). The inverse solution was computed by brainstorm built-in minimum norm estimation algorithm applying with the default settings ("kernel only" as the output mode, 3 as the signal-to-noise ratio, the source orientation constrained to perpendicular to the cortical surface, the depth weighting restricting source locations to the cortical surface and the whitening PCA). A noise covariance matrix, necessary to control noise effects on the solution (Bouhamidi, & Jbilou, 2007) was calculated over –250 to –150 baseline interval (Dale et al., 2000).

The individual source maps were projected to the cortical surface of the Montreal Neurological Institute brain template (MNI-Colin27). Differences in source activation between verb generation and silent reading were tested via paired t-tests under significance level of $p < 0.05$, uncorrected.

Results

Behavioral results

The average response time in overt verb generation task was 1.56 sec (SD = 0.2). Mean accuracy was 7% (SD = 3.9). The responses which were considered incorrect were removed from the subsequent analysis.

General time course of noun cue recognition

As shown in butterfly plots for VG and SR conditions (Figure 1B), the general pattern of the response to the written noun roughly coincided in both tasks, presumably reflecting the common process of word recognition. The noun presentation elicited the evoked response with two early narrow peaks around 100 and 140 ms after word onset followed by broader components at 200 and 400 ms. Based on extensive literature on visual word recognition (for recent reviews see Carreiras, Armstrong, Perea, & Frost, 2014; Grainger, & Holcomb, 2009) we identified the ob-

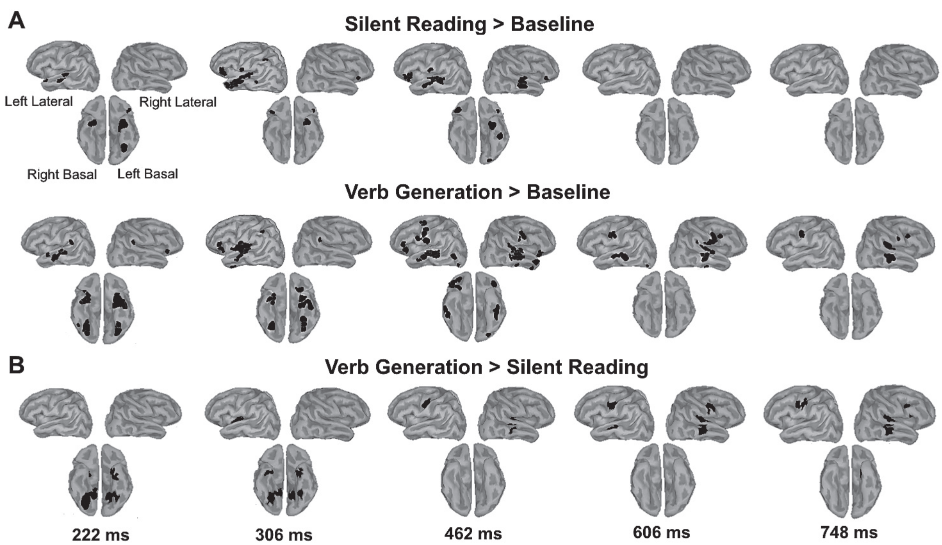


Figure 1. The brain response to the written nouns in verb generation (VG) and silent reading (SR) tasks: sensor-level analysis. (A) Three projections (SPM glass image) show the sensor array from above (transverse), the right (sagittal), and the back (coronal). A — anterior, P — posterior, L-left and R-right parts of the array. Areas in black correspond to spatial clusters with significant sensor-level differences in ERF between VG and SR task (paired t-test, $p < 0.05$, FDR-corrected). All the clusters reflect greater response under VG versus SR condition and occur within four time windows of 191–227 ms, 306–340 ms, 462–619 ms and 676–891 ms after the noun onset. The local peaks are reported as small black circles. Note, that no clusters with opposite direction of the effect were found. (B) Butterfly plots of MEG evoked waveforms from 306 MEG channels. Strength of magnetic fields is represented in femto-Tesla (fT). Zero point denotes the onset of the noun cue. The increase of the response around zero is presumably related to the stronger attention allocation to the fixation cross under VG condition. Shaded rectangles denote the time windows with significant VG-SR difference. The gray lines indicate the time points of the local peaks of SPM clusters

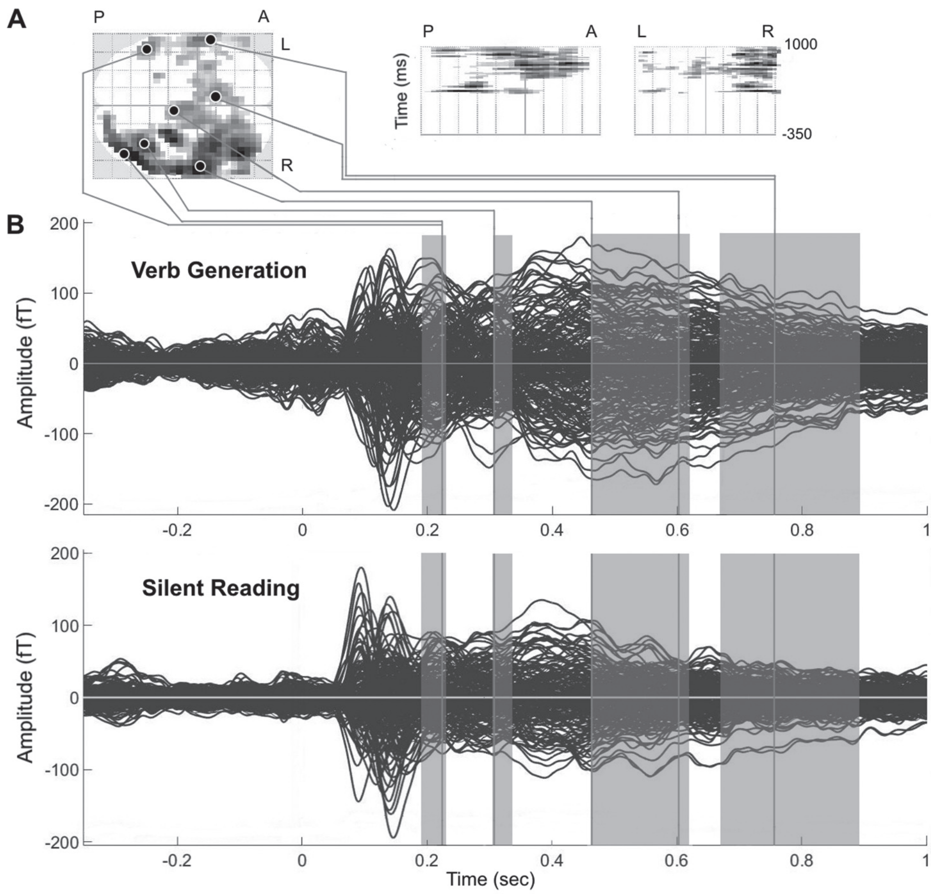


Figure 2. Reconstructed cortical activation that contributes to significant SR-VG differences. The time points correspond to the temporal peaks of sensor-level significant difference. (A) Cortical response evoked by visually presented noun in verb generation and silent reading task. Note that noun-evoked activation progressed along the posterior-anterior axis from posterior sensory regions to more anterior multimodal association areas. (B) The reconstruction of cortical sources displaying greater response in VG versus SR task. All images were thresholded using a voxel-wise statistical threshold of $p<0.05$ with cluster size more than 10 voxels

served ERF deflections as MEG counterparts of P100, N170, N200 and N400 components established in EEG studies (e.g. Hauk, & Pulvermüller, 2004; Holcomb, & Grainger, 2006; Kutas, & Federmeier, 2011; Maurer, Brandeis, & McCandliss, 2005). MNE source estimation revealed that the response emerged around 100 ms at the occipital cortex bilaterally, then shifted toward the inferior occipitotemporal cortex at 140 ms with a prominent left-hemispheric preponderance (the results are presented elsewhere (Butorina et al, submitted)). By 190 ms the activation reached the anterior part of temporal lobes (ATL) bilaterally. During 200-300 ms the response engaged also the lateral surface of the left hemisphere, namely ventrolateral prefrontal cortex (VLPFC), superior and middle temporal gyri (STG, MTG) and the perisylvian region (Figure 2A). Around 460 ms the response in the temporal cortex became bilateral, and was accompanied by left-lateralized activity in ventral pericentral region and VLPFC. After 600 ms the response in SR decreased rapidly in contrast to VG where bilateral temporal and pericentral activity was present for another two hundreds milliseconds — up to 800 ms post-stimulus. The spatial-temporal patterns observed under SR and VG conditions are in good correspondence with features of the brain response described for word recognition (e.g. Dien, 2009; Grainger, & Holcomb, 2009; Pulvermüller, Shtyrov, & Hauk, 2009; Salmelin, 2007), thus, confirming that both tasks consecutively activated the same brain areas associated with the written word processing.

Task effects on noun cue recognition
Sensor-level analysis

Figure 1A presents the spatial-temporal clusters of sensor-level SR-VG differences that survived cluster-level FDR correction with a threshold at $p < .05$. The evoked response to a written noun word was stronger in VG compared to SR task for all the significant clusters, while no clusters demonstrating the opposite direction of the task effect were found. As shown in the butterfly plots (Figure 1B) two earliest components at 100 and 140 ms (P100m, N170m) remained unaffected by the task demands. The initial task-related difference in the response strength appeared at time window of N200m component — at 191–227 ms — and was concentrated over the posterior boundary sensors on the right side of the array ($p < .0001$, FDR-corr.) and over the lateral sensors on the left ($p < .0001$, FDR-corr.). The following cluster of differential response emerged at 306–340 ms over the postetior-lateral sensors of the right hemisphere ($p < .0001$, FDR-corr.). After that, the peak of differential activity shifted to the anterior half of the sensor array. Time window from 460 to 900 ms post-stimulus was dominated by the widespread clusters over the right anterior and lateral sensors at 462–619 and 676–891 ms time window ($p < .0001$, FDR-corr.) while we also detected the symmetrical but smaller clusters of SR-VG difference in the left-hemispheric part of the sensor array. Thus, the task effects on written word processing were highly reliable and were confined to four consecutive time windows with gradually increasing duration.

Source-level analysis

To determine the cortical areas that contribute into significant SR-VG differences in the brain response at the sensor-level we used temporal clusters as a mask defining the time windows of interest for source-level analysis. Given that the source-space analysis was guided by FDR-corrected sensor-level results, the statistical threshold in the source-space was defined at $p < .05$ (peak-level, uncorrected) with cluster size more than 10 adjunct voxels.

The early time window — 191–227 ms post-stimulus — was characterized by the differential response at the regions on the basal surface of the occipitotemporal lobes (Figure 2B). The activity in the left ATL and the left inferior occipitotemporal cortex was presented under both conditions but was stronger in VG than in SR, while the right occipitotemporal region was recruited into the response only under VG condition. During following time window at 306-340 ms the task effect in the ATL and inferior occipitotemporal cortex was bilateral and also engaged the left transverse gyrus comprising primary auditory cortex. At 462 ms the difference in the response was observed in the right STG and MTG and in the left ventral sensorimotor cortex - in the inferior region of left precentral and postcentral gyri. The differential activation of these regions spread to homotopic areas of both hemispheres, reached its maximum at 600 ms, and then sustained for another 200 ms up to the end of measurable brain response.

Discussion

Here we present MEG evidence that task demands penetrate into early (200-300 ms) and late (500-800 ms) stages of written word processing. We varied task demands for silent reading of a written noun word by imposing an instruction either to perform no further action (silent reading — SR) or to name a related verb afterwards (verb generation — VG). The long verb production time (1.5 s on average) in the latter case precluded the possibility that any changes in the processing of a written noun within 100–800 ms after its presentation onset was simply elicited by a preparation of motor response.

Our data indicates that top–down modulation affects relatively early processes in visual word recognition. More difficult task demands enhanced brain response at the stage of word form processing (Figure 2), while spared the low-level visual processing of stimulus features, i.e. contrast, figure-ground segregation etc. A lack of task effects on average word activation at the latency of the P100m and N100m component (Figure 1) signified that our tasks were similar with respect to the visual attention, which is known to increase these ERP components (e.g. Hillyard & Anllo-Vento, 1998). The visual word form area (VWFA) located in the left inferior occipitotemporal cortex and the left temporal pole were the first brain areas to show the elevated response to VG as compared to SR condition at 190–230 ms (Figure 2). With 80 ms delay the similar differential activation was observed in the homotopic areas of the right hemisphere (Figure 2). While VWFA is thought to recognize a letter string as a word form stored in long-term memory (Cohen et al., 2000; Dehaene, & Cohen, 2011), bilateral regions of temporal pole have been assigned a role of a semantic hub linking word forms with distributed representations of the same word in different sensory modalities (Binney, Embleton, Jefferies, Parker, & Lambon Ralph, 2010; Patterson, Nestor, & Rogers, 2007). Greater engagement of the temporal pole within 200–250 ms window in VG task may promote early semantic analysis of "word meaningfulness", which, according to EEG reading research, also peaks around 250 ms after the word onset (for a review see Martín-Loeches, 2007).

Moreover, our findings clearly contradict the claims that the role of VWFA is "strictly visual and prelexical" (Dehaene, & Cohen, 2011). Our observation of effects "what-I-will-do-next-with this word" on activation in VWFA and more anterior part of ventral temporal lobe lends a firm support to an interactive view on the role of VWFA and other areas of ventral visual stream in word processing (Price, & Devlin, 2011).

In this respect, we extend and substantiate the previously existing literature, which implied VWFA activity to be sensitive to semantic or lexical decision directly related to a written word (Chen, Davis, Pulvermüller, & Hauk, 2015), and consequently was not immune to the unspecific modulatory effect of "word targetness". Our data revealed a greater involvement of VWFA and its right hemispheric counterpart even in the case when the presented word by itself did not represent the target of the task. We speculate that a participant's need to further proceed with the retrieved information regarding the written word intensifies and deepens visual processing within the first 200-300 msec after a word presentation through integration of the task demands with the processing of immediate visual input.

The shift of activation from basal posterior temporal to lateral anterior areas of temporal and frontal lobes (Figure 2) at the latencies of N400m was common for both SR and VG tasks and indicated that the crucial point of the word recognition process in each task was the retrieval of word semantics. It may therefore surprise that we did not find any reliable task effects within the time window of "semantic" N400 component (Figure 2). Our VG task required participants to choose the verb associated with the presented noun, thus, encouraging the retrieval of noun's semantic features related to actions, and the difficulty of such retrieval did affect the "semantic brain activity" within N400 time window in our previous analysis (Butorina et al., submitted). One possibility is that the existing task effect did not survive the rigorous statistical corrections performed in the current study. Certainly, further research is needed to clarify the origin of this puzzling result.

In addition to the early task modulations, it is also striking that we found a robust but late (500 ms after the noun onset) and temporally protracted (500 - 800 ms) effect of task demands on the activity of lateral temporal regions and to a lesser extent of ventral posterior frontal regions. Interestingly, this effect was more prominent in the right hemisphere. This finding is not supported by the previous literature that similarly to our study used non-invasive MEG recording of human brain activity. However, our results are fully consistent with those obtained with subdural electrodes implanted in various brain regions of pre-surgical epilepsy patients. In this study the patients performed different variants of an overt word repetition task thoughtfully designed to check a hypothesis on the role of sensorimotor activations in speech perception (Cogan et al., 2014). Critically, the ECoG study observed that word presentation elicited a robust bilateral activation in the middle temporal gyrus, superior temporal gyrus, somatosensory, motor and premotor cortex, as well as in supramarginal gyrus and inferior frontal gyrus, i.e. roughly in the same cortical regions and at the same latency as in the current MEG study. Manipulating various task demands the authors proved that this bilateral activation represents sub-vocal sensorimotor transformations that is specific for speech and unifies perception- and production-based representations to facilitate access to higher order language functions. Given the evidence, our findings on enhanced late response in the same areas during even more demanding task may reflect enhanced sensorimotor transformation for the already perceived noun in case of a need for further decision making.

A long-standing notion is that motor system may contribute to speech perception by internally emulating sensory consequences of articulatory gestures (Liberman & Mattingly, 1985; Liberman & Wahlen, 2000). This does not mean that this mechanism is necessary to understand speech. Rather it serves as an assisting device when word-related decision making is facing difficult circumstances. Previous studies have shown that consulting internal sensorimotor models ameliorates speech perception when the speech input is ambiguous and/or noisy (Meister et al., 2007; Möttönen and Watkins, 2009). Our study reveals the same mechanism coming into play when there is a need to maintain and manipulate with perceived written noun in the working memory in order to produce its verb associates. Reverberating sensorimotor loops related to inner phonemic emulation of the presented word has been linked to phonological working memory by other authors (Hickok, & Poeppel, 2007; Buchsbaum et al., 2011). However, our results demonstrates for the first time that an intensified engagement of ventral motor cortex produced by a need of deep word processing is automatic and does not happened consciously as a sub-vocal rehearsal.

Intriguingly, higher task demands in our study seem to increasingly recruit the right hemisphere networks in such transformative activity. In accord with Cogan's et al hypothesis our MEG data obtained in typical participants argue against the prevailing dogma that dorsal stream sensorimotor functions are highly left lateralized. Thus, the current study contributes into the growing evidence from lesion, imaging, and electrophysiological data demonstrating convincingly the complex lateralization patterns for different language operations (for review, see Poeppel, 2014).

Conclusion

Our results suggest that a remote goal plays a pivotal role in enhanced recruitment of cortical structures underlying orthographic, semantic and sensorimotor dimensions of written word perception from the early processing stages. They also show that passive speech perception induces activation of brain areas involved in speech production. The increased recruitment of these areas in a more demanding task could reflect an automatic mapping of phonemes onto the articulatory motor programs – the process involved in covert imitative mechanisms or internal speech, which might, in turn, improve comprehension of the percept. During silent reading a need in deep processing of linguistic input may play a central role in linking speech perception with speech production, consistent with theories that emphasize the integration of sensory and motor representations in understanding speech (Hickok, & Poeppel, 2000; Scott, & Wise, 2004). Surprisingly, we found that in order to fulfil a more challenging goal the brain progressively engaged the resources of the right-hemisphere throughout all stages of silent reading. This conclusion fits well with mounting evidence on the role of right hemisphere in speech perception and higher-order cognition (Velichkovsky, Krotkova, Sharaev, & Ushakov, 2017).

Acknowledgments

This study was in part supported by the Russian Science Foundation, grant RScF 14-28-00234. We wish to thank B.M. Velichkovsky for discussion of results.

References

Balota, D.A., Black, S.R., & Cheney, M. (1992). Automatic and attentional priming in young and older adults: Reevaluation of the two-process model. *Journal of Experimental Psychology: Human Perception and Performance, 18*(2), 485–502. doi: 10.1037/0096-1523.18.2.485

Balota, D.A., Burgess, G.C., Cortese, M.J., & Adams, D.R. (2002). The word-frequency mirror effect in young, old, and early-stage Alzheimer's disease: Evidence for two processes in episodic recognition performance. *Journal of Memory and Language, 46*(1), 199–226. https://doi.org/10.1006/jmla.2001.2803

Balota, D.A., & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 340–357. doi: 10.1037/0096-1523.10.3.340

Balota, D.A., Cortese, M.J., Sergent-Marshall, S.D., Spieler, D.H., & Yap, M.J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283–316. doi: 10.1037/0096-3445.133.2.283

Balota, D.A., & Yap, M.J. (2006). Attentional control and the flexible lexical processor: Explorations of the magic moment of word recognition. *From Inkmarks to Ideas: Current Issues in Lexical Processing*, 229.

Bentin, S., Kutas, M., & Hillyard, S.A. (1993). Electrophysiological evidence for task effects on semantic priming in auditory word processing. *Psychophysiology*, *30*(2), 161–169. doi: 10.1111/j.1469-8986.1993.tb01729.x

Bentin, S., Kutas, M., & Hillyard, S.A. (1995). Semantic processing and memory for attended and unattended words in dichotic listening: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 54–67. doi: 10.1037/0096-1523.21.1.54

Besner, D. (2001). The myth of ballistic processing: Evidence from Stroop's paradigm. *Psychonomic Bulletin & Review*, *8*(2), 324–330. doi: 10.3758/BF03196168

Besner, D., Stolz, J.A., & Boutilier, C. (1997). The stroop effect and the myth of automaticity. *Psychonomic Bulletin & Review*, *4*(2), 221–225. doi: 10.3758/BF03209396

Binney, R.J., Embleton, K.V., Jefferies, E., Parker, G.J.M., & Lambon Ralph, M.A. (2010). The ventral and inferolateral aspects of the anterior temporal lobe are crucial in semantic memory: evidence from a novel direct comparison of distortion-corrected fMRI, rTMS, and semantic dementia. *Cerebral Cortex*, *20*(11), 2728-2738. doi: 10.1093/cercor/bhq019

Bodner, G.E., & Masson, M.E.J. (2004). Beyond binary judgments: Prime validity modulates masked repetition priming in the naming task. *Memory & Cognition*, *32*(1), 1–11. doi: 10.3758/BF03195815

Bouhamidi, A., & Jbilou, K. (2007). Sylvester Tikhonov-regularization methods in image restoration. *Journal of Computational and Applied Mathematics*, *206*(1), 86–98. doi: 10.1016/j.cam.2006.05.028

Buchsbaum, B.R., Baldo, J., Okada, K., Berman, K.F., Dronkers, N., D'Esposito, M., & Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory–an aggregate analysis of lesion and fMRI data. *Brain and Language*, *119*(3), 119–128. doi: 10.1016/j.bandl.2010.12.001

Butorina A.V., Pavlova A.A., Nikolaeva A.Y., Prokofyev A.O., Bondarev D.P., & Stroganova T.A. (2017). Simultaneous processing of noun cue and to-be-produced verb in verb generation task: Electromagnetic evidence. *Frontiers in Human Neuroscience, 11*, 279. doi: 10.3389/fnhum.2017.00279

Carr, T.H., & Dagenbach, D. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *16*(2), 341–350. doi: 10.1037/0278-7393.16.2.341

Carreiras, M., Armstrong, B.C., Perea, M., & Frost, R. (2014). The what, when, where, and how of visual word recognition. *Trends in Cognitive Sciences*, *18*(2), 90–98. doi: 10.1016/j.tics.2013.11.005

Chen, Y., Davis, M. H., Pulvermüller, F., & Hauk, O. (2015). Early visual word processing is flexible: Evidence from spatiotemporal brain dynamics. *Journal of Cognitive Neuroscience, 27*(9), 1738–1751. doi: 10.1162/jocn_a_00815

Cogan, G. B., Thesen, T., Carlson, C., Doyle, W., Devinsky, O., & Pesaran, B. (2014). Sensory–motor transformations for speech occur bilaterally. *Nature*, *507*(7490), 94–98. doi: 10.1038/nature12935

Cohen, L., Dehaene, S., Naccache, L., Lehericy, S., Dehaene-Lambertz, G., Henaff, M. A., & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*. doi: 10.1093/brain/123.2.291

Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., & Halgren, E. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, *26*(1), 55–67. doi: 10.1016/S0896-6273-(00)81138-1

Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262. doi: 10.1016/j.tics.2011.04.003

Dien, J. (2009). The neurocognitive basis of reading single words as seen through early latency ERPs: A model of converging pathways. *Biological Psychology*, *80*(1), 10–22. doi: 10.1016/j.biopsycho.2008.04.013

Ferrand, L., Brysbaert, M., Keuleers, E., New, B., Bonin, P., Méot, A., … Pallier, C. (2011). Comparing word processing times in naming, lexical decision, and progressive demasking: Evidence from Chronolex. *Frontiers in Psychology*, *2*, 306. doi: 10.3389/fpsyg.2011.00306

Fodor, J.A. (1983). *The modularity of mind: An essay on faculty psychology*. MIT press.

Forster, K.I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 680–698. doi: 10.1037/0278-7393.10.4.680

Grainger, J., & Holcomb, P. (2009). Watching the word go by: On the time-course of component processes in visual word recognition. *Language and Linguistics Compass*, *3*(1), 128–156. doi: 10.1111/j.1749-818x.2008.00121.x

Hämäläinen, M.S., & Ilmoniemi, R.J. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Medical {&} Biological Engineering {&} Computing*, *32*(1), 35–42. doi: 10.1007/BF02512476

Hauk, O., & Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology*, *115*(5), 1090–1103. doi: 10.1016/j.clinph.2003.12.020

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402. doi: 10.1038/nrn2113

Hillyard, S.A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences of the United States of America*, *95*(3), 781–787. doi: 10.1073/pnas.95.3.781

Holcomb, P.J., & Grainger, J. (2006). On the time course of visual word recognition: An event-related potential investigation using masked repetition priming. *Journal of Cognitive Neuroscience*, *18*(10), 1631–1643. doi: 10.1162/jocn.2006.18.10.1631

Holland, P.W., & Welsch, R.E. (1977). Robust regression using iteratively reweighted least-squares. *Communications in Statistics — Theory and Methods*, *6*(9), 813–827. doi: 10.1080/03610927708827533

Huang, M.X., Mosher, J.C., & Leahy, R.M. (1999). A sensor-weighted overlapping-sphere head model and exhaustive head model comparison for MEG. *Physics in Medicine and Biology*, *44*(2), 423. Retrieved from http://stacks.iop.org/0031-9155/44/i=2/a=010

Kawamoto, A.H., Kello, C.T., Higareda, I., & Vu, J.V.Q. (1999). Parallel processing and initial phoneme criterion in naming words: Evidence from frequency effects on onset and rime duration. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 362–381. doi: 10.1037/0278-7393.25.2.362

Kilner, J.M., & Friston, K.J. (2010). Topological inference for EEG and MEG. *The Annals of Applied Statistics*, *4*(3), 1272–1290. doi: 10.1214/10-AOAS337

Kissler, J., Herbert, C., Winkler, I., & Junghofer, M. (2009). Emotion and attention in visual word processing — An ERP study. *Biological Psychology*, *80*(1), 75–83. doi: 10.1016/j.biopsycho.2008.03.004

Kutas, M., & Federmeier, K.D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). *Annual Review of Psychology*, *62*, 621. doi: 10.1146/annurev.psych.093008.131123

Liberman, A.M., & Mattingly, I.G. (1985). The motor theory of speech perception revised. *Cognition*, *21*(1), 1–36. doi: 10.1016/0010-0277(85)90021-6

Liberman, A.M., & Whalen, D.H. (2000). On the relation of speech to language. *Trends in cognitive sciences*, *4*(5), 187–196. doi: 10.1016/S1364-6613(00)01471-6

Lyashevskaya, O.N. & Sharov, S.A. (2009) *Chastotnyy slovar'sovremennogo russkogo yazyka (na materialakh natsional'nogo korpusa russkogo yazyka).* [The frequency dictionary of modern Russian language (based on the Russian National Corpus)]. Moscow: Azbukovnik.

MacLeod, C.M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163–203. doi: 10.1037/0033-2909.109.2.163

MacLeod, C. M. (2005). The Stroop Task in cognitive research. In *Cognitive methods and their application to clinical research.* (pp. 17–40). Washington: American Psychological Association. doi: 10.1037/10870-002

Mahé, G., Zesiger, P., & Laganaro, M. (2015). Beyond the initial 140ms, lexical decision and reading aloud are different tasks: An ERP study with topographic analysis. *NeuroImage*, *122*, 65–72. doi: 10.1016/j.neuroimage.2015.07.080

Marcel, A.J. (1983). Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology*, *15*(2), 238–300. doi: 10.1016/0010-0285(83)90010-5

Martín-Loeches, M. (2007). The gate for reading: Reflections on the recognition potential. *Brain Research Reviews*, *53*(1), 89–97. doi: 10.1016/j.brainresrev.2006.07.001

Maurer, U., Brandeis, D., & McCandliss, B.D. (2005). Fast, visual specialization for reading in English revealed by the topography of the N170 ERP response. *Behavioral and Brain Functions*, *1*(1), 13. doi: 10.1186/1744-9081-1-13

Meister, I.G., Wilson, S.M., Deblieck, C., Wu, A.D., & Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Current Biology*, *17*(19), 1692-1696. doi: 10.1016/j.cub.2007.08.064

Möttönen, R., & Watkins, K.E. (2009). Motor representations of articulators contribute to categorical perception of speech sounds. *Journal of Neuroscience*, *29*(31), 9819–9825. doi: 10.1523/JNEUROSCI.6018-08.2009

Neely, J.H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. *Basic Processes in Reading: Visual Word Recognition*, *11*, 264–336.

Neely, J.H., & Kahan, T.A. (2001). Is semantic activation automatic? A critical re-evaluation. In *The nature of remembering: Essays in honor of Robert G. Crowder.* (pp. 69–93). American Psychological Association.

Patterson, K., Nestor, P.J., & Rogers, T.T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Review of Neuroscience*, *8*(12), 976–987. doi: 10.1038/nrn2277

Poeppel, D. (2014). The neuroanatomic and neurophysiological infrastructure for speech and language. *Current Opinion in Neurobiology*, *28*, 142–149. doi: 10.1016/j.conb.2014.07.005

Posner, M.I., & Snyder, C.R.R. (1975). Facilitation and inhibition in the processing of signals. *Attention and Performance V*, 669–682.

Price, C.J., & Devlin, J.T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. *Trends in Cognitive Sciences*, *15*(6), 246–253. doi: 10.1016/j.tics.2011.04.001

Pulvermüller, F., Shtyrov, Y., & Hauk, O. (2009). Understanding in an instant: Neurophysiological evidence for mechanistic language circuits in the brain. *Brain and Language*, *110*(2), 81–94. doi: 10.1016/j.bandl.2008.12.001

Ruz, M., & Nobre, A.C. (2008). Attention modulates initial stages of visual word processing. *Journal of Cognitive Neuroscience*, *20*(9), 1727–1736. doi: 10.1162/jocn.2008.20119

Salmelin, R. (2007). Clinical neurophysiology of language: The MEG approach. *Clinical Neurophysiology*, *118*(2), 237–254. doi: 10.1016/j.clinph.2006.07.316

Schilling, H.E H., Rayner, K., & Chumbley, J.I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, *26*(6), 1270–1281. doi: 10.3758/BF03201199

Strijkers, K., Yum, Y.N., Grainger, J., & Holcomb, P.J. (2011). Early goal-directed top-down influences in the production of speech. *Frontiers in Psychology*, *2*, 371. doi: 10.3389/fpsyg.2011.00371

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. doi: 10.1037/h0054651

Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., & Leahy, R.M. (2011). Brainstorm: A user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience*, *2011*, 8. doi: 10.1155/2011/879716

Taulu, S., Simola, J., & Kajola, M. (2005). Applications of the signal space separation method. *IEEE Transactions on Signal Processing*. doi: 10.1109/TSP.2005.853302

Velichkovsky, B.M. (2006). *Cognitivnaya nauka: Osnovy psikhologii poznaniya* [Cognitive science: Foundations of epistemic psychology]. (Vol. 1). Moscow: Smysl / Academia.

Velichkovsky, B.M., Krotkova, O.A., Sharaev, M.G., & Ushakov, V.L. (2017). In search of the "I": Neuropsychology of lateralized thinking meets Dynamic Causal Modeling. *Psychology in Russia: State of the Art, 10*(3), 7–27. doi: 10.11621/pir.2017.0301

Yap, M.J., Pexman, P.M., Wellsby, M., Hargreaves, I.S., & Huff, M.J. (2012). An abundance of riches: Cross-task comparisons of semantic richness effects in visual word recognition. *Frontiers in Human Neuroscience*, *6*, 72. doi: 10.3389/fnhum.2012.00072

Original manuscript received May 29, 2017
Revised manuscript accepted June 13, 2017
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society · Lomonosov Moscow State University

Cognitive aspects of human motor activity: Contribution of right hemisphere and cerebellum

Aleksei S. Sedov[a,b*], Valentin A. Popov[a], Veronika I. Filyushkina[a], Ulia N. Semenova[a], Viacheslav A. Orlov[c], Boris M. Velichkovsky[b,c,e], Vadim L. Ushakov[c,d]

[a] *Semenov Institute of Chemical Physics, Russian Academy of Sciences, Moscow, Russia*
[b] *Moscow Institute of Physics and Technology, Moscow, Russia*
[c] *Kurchatov Institute National Research Centre, Moscow, Russia*
[d] *Lomonosov Moscow State University, Moscow, Russia*
[e] *Russian State University for the Humanities, Moscow, Russia*

* Corresponding author. E-mail: sedov.chph@yandex.ru

Background. Concepts of movement and action are not completely synonymous, but what distinguishes one from the other? Movement may be defined as stimulus-driven motor acts, while action implies realization of a specific motor goal, essential for cognitively driven behavior. Although recent clinical and neuroimaging studies have revealed some areas of the brain that mediate cognitive aspects of human motor behavior, the identification of the basic neural circuit underlying the interaction between cognitive and motor functions remains a challenge for neurophysiology and psychology.

Objective. In the current study, we used functional magnetic resonance imaging (fMRI) to investigate elementary cognitive aspects of human motor behavior.

Design. Twenty healthy right-handed volunteers were asked to perform stimulus-driven and goal-directed movements by clenching the right hand into a fist (7 times). The cognitive component lay in anticipation of simple stimuli signals. In order to disentangle the purely motor component of stimulus-driven movements, we used the event-related (ER) paradigm. fMRI was performed on a 3 Tesla Siemens Magnetom Verio MR-scanner with 32-channel head coil.

Results. We have shown differences in the localization of brain activity depending on the involvement of cognitive functions. These differences testify to the role of the cerebellum and the right hemisphere in motor cognition. In particular, our results suggest that right associative cortical areas, together with the right posterolateral cerebellum (Crus I and lobule VI) and basal ganglia, define cognitive control of motor activity, promoting a shift from a stimulus-driven to a goal-directed mode.

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0314
http://psychologyinrussia.com

Conclusion. These results, along with recent data from research on cerebro-cerebellar circuitry, redefine the scope of tasks for exploring the contribution of the cerebellum to diverse aspects of human motor behavior and cognition.

Keywords: action, movement, fMRI, lateralization, motor behavior, voluntary movement, cognition, cortex, cerebellum, basal ganglia

Introduction

Motor acts are not just made; they are thought out, planned, organized, and learned. These all require the involvement of various integrated cognitive functions, allowing for a successful performance. The neural substrate of cognitive aspects of movement remains a matter of debate. The classic model of motor control describes a pyramidal system as the main executive part and an extrapyramidal system that controls smoothness and precision, while in addition providing feedback. Human motor activity is carried out by activating a number of cortical structures: the primary motor (M1) and somatosensory (PSC) cortex, the premotor (PMC) cortex, and the supplementary motor area (SMA) (Grefkes, Eickhoff, Nowak, Dafotakis, & Fink, 2008; Wu, Kansaku, & Hallett, 2004). Patterns of activation were also observed in the cingulate gyrus, several parts of the cerebellum, thalamus, and basal ganglia (Wardman, Gandevia, & Colebatch, 2014).

Investigation of brain activity during motor performance has revealed other non-motor structures testifying to motor–cognitive interactions (Gentsch, Weber, Synofzik, Vosgerau, & Schütz-Bosbach, 2016; Strick, Dum, & Fiez, 2009). Parietal and frontal areas have been found to be part of dorsal and ventral attention systems and accordingly play a role in detection of behaviorally relevant sensory events, therefore influencing the motor response (Corbetta, & Shulman, 2002). Insula – also a region of interest in motor cognition studies – have been hypothesized to be an area of convergence for these attentional systems (Nelson et al, 2009). The basal ganglia may support a basic attentional mechanism to bind input to output in the executive forebrain, which provides the link between voluntary effort and the operation of a sequence of motor programs or thoughts (Wu, Kansaku, & Hallett, 2004).

Recent studies have proposed a habitual and goal-directed control impairment model for movement disorders such as Parkinson's disease and for some mental disorders (Jahanshahi, Obeso, Rothwell, & Obeso, 2015; Redgrave, Rodriguez, Smith, Rodriguez-Oroz, Lehericy, Bergman, & Obeso, 2010). Along with numerous neuroimaging and psychological findings, they suggest that while most actions are expressed in movement, they tend to involve higher-level processes such as sensory integration, motor planning, or decision making. There is an intellectual aspect of movement control, which is obligatory for an adaptive action.

Experiments with motor, cognitive, and motor/cognitive tests and procedures have revealed a vast array of brain areas responding, depending upon the task (Behroozmand et al., 2015; Von der Gablentz, Tempelmann, Münte, & Heldmann 2015). Cerebellar activation of diverse types could be seen in a variety of studies (Stoodley, & Schmahmann, 2009) examining its role in motor behavior control. The classical symptoms of cerebellar lesion – such as ataxia, negative Rhomberg's test, and vertigo – all involve coordination of voluntary movements, posture, and equilibrium. After Schmahmann's and Sherman's report on cerebellar cognitive affective syndrome, the notion of cerebellar functions, mainly concerned with control and coordination of motor activity, required broadening (Schmahmann, & Sherman, 1998).

The variety of behavioral deficits including executive, visual–spatial, linguistic, and emotional impairment suggests a constellation of circuits linking the cerebellum with vast brain areas of different functional modality. There is now no doubt that a significant, albeit not yet specified, part of the cerebellar output projects to non-motor areas (Allen, & Tsukahara, 1974; Anand, Malhotra, Singh, & Dua, 1959). Anatomical evidence that the cerebellum exerts an influence over non-motor regions of the cerebral cortex is complemented by data from neuroimaging and neuropsychology (Appollonio, Grafman, Schwartz, Massaquoi, & Hallett, 1998; Botez-Marquard, Léveillé, & Botez, 1994). These lines of research have provided compelling evidence that the cerebellum plays a functionally important role in human cognition. In this light, we planned the present study to include systematic observations of cerebral and cerebellar activation during goal-directed movements.

Method

Twenty right-handed healthy volunteers (11M, 9F), with a mean age of 22 ± 3 years, participated in this study. All subjects were carefully instructed about MR investigation features and conditions and were included only after signing an informed consent.

Brain imaging was performed on a 3 Tesla SIEMENS Magnetom Verio MR-scanner with 32-channel head coil. Head motion was reduced by a belt around the subject's head. Subjects lay supine in the MR scanner with a response device fixed to their right hand. The protocol included: 1) T1-weighted sagittal three-dimensional magnetization-prepared rapid gradient echo sequence (176 slices, TR = 1470 ms, TE = 1.76 ms, voxel size 1x1x1 mm) for anatomical data and 2) T2 EPI echo planar sequence (42 slices, TR = 2000 ms, TE = 44 ms, voxel size 1.5 × 1.5 × 2.6 mm) for functional images. The ultrafast fMRI sequence was obtained from the University of Minnesota Center for Magnetic Resonance Research. Also we received data that contain options for reducing the spatial distortion of EPI images.

We employed two different paradigms. In order to disentangle the purely motor component of stimulus-driven movements we used the event-related (ER) paradigm. Volunteers were asked to clench their right hand into a fist (7 times) in response to verbal commands. We studied brain activity only during the movement itself, with an action period of 1 second each.

The second paradigm was used to study the cognitive aspects of goal-directed movements. This block-design paradigm consisted of 7 alternating rest and action periods of 30 seconds each. During the action period, evenly played beeping was introduced to the volunteers, who were asked to clench their hand anticipating the beep, which was followed by another in 1.5 seconds. These movements were considered goal-directed, involving a more complex, comprehensive behavior composed of both motor and cognitive aspects such as attention and time appreciation, in comparison with stimulus-driven movements.

The Matlab (MathWorks) free access SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) package was used for parametric mapping of anatomical and functional data. Preliminary processing included DICOM files transferring, fixing anatomical coordinates to AC-PC line, image correction considering head tilt, matching anatomy with activity clusters and their equalization, and Gaussian smoothing of the data. A design matrix was then created. Moments when the stimulus would be presented, along with action duration, were set. Then parametric statistical mapping of the brain areas was used with the common linear model GLM.

The confidence interval for individual and group analysis was chosen according to T-criterion ($p < 0.05$), considering multiple comparison test (FWE) and topological adjustment FDR ($q<0.05$). Cluster localization and analysis was performed with SPM packages Anatomy Toolbox (Eickhoff et al., 2005) and WFU PickAtlas (Maldjian, Laurienti, Kraft, & Burdette, 2003).

Results

In the current study, we analyzed fMRI data obtained from healthy right-handed participants who were asked to perform voluntary movements. We observed activations in the five clusters: 1) contralateral pre- and postcentral gyri (738 voxels); 2) contralateral insula (31 voxels); 3) supplementary motor area (SMA) (56 voxels); 4) ventral thalamic nuclei (VPL, VPM) (22 voxels); 5) ipsilateral cerebellum lobules IV and V (134 voxels) projection as well as lobules VI (36 voxels) projection (Figure 1).

Significant activation in goal-directed movements compared to baseline was observed in total within 19 clusters (1,271 voxels) (Fig. 2, Table 1). Motor cortical clusters consisted of 264 voxels and included the primary motor cortex (M1) and the somatosensory cortex (PSC). Cortical regions also included bilateral supplementary motor area (SMA), ipsilateral inferior parietal lobule, supramarginal gyrus, superior frontal gyrus, and frontal operculum. In subcortical structures, several clusters (167 voxels) in total were observed. The most significant one was in

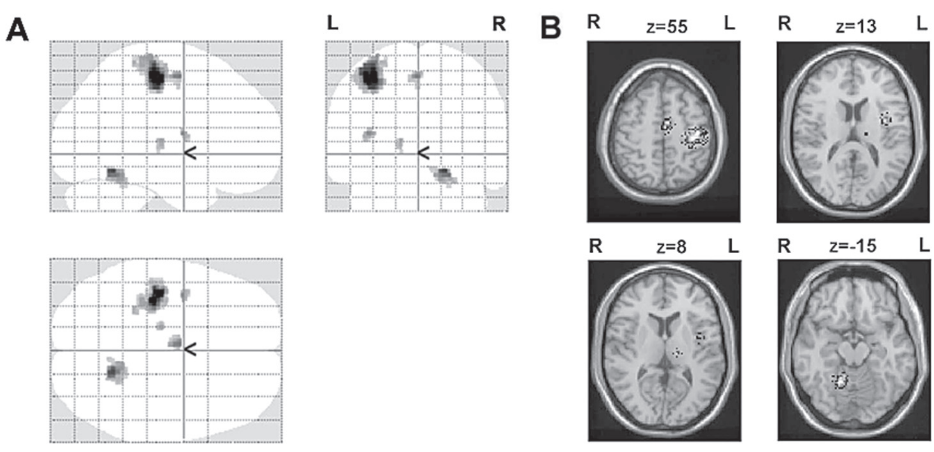


Figure 1. Brain areas activated during stimulus-driven movement. A — statistical parametric maps of activated areas, B — activated areas imposed on averaged brain

ventral thalamic nuclei (51 voxels). Basal ganglia were presented with contralateral anterior striatum areas and ipsilateral putamen and pallidal areas (18 voxels). The last group of clusters that we obtained using this paradigm was bilateral cerebellar (1013 voxels). The largest of them (707 voxels) involved the ipsilateral IV, V (335 voxels) and VI (161 voxels) lobules, along with vermic lobule IV, V (153 voxels), VI (34 voxels), and VIII (3 voxels). In the contralateral cerebellar hemisphere cluster (288 voxels) in VI lobule and Crus I was activated.

Table 1. Localization of brain areas activated during goal-directed movement.

No.	Cluster size	MNI coordinates			Activated areas			T-value
		X	Y	Z	Side	Structure	Size	
1	264	–40.5	–22	57.4	L	Precentral gyrus	196	6.8
					L	Postcentral gyrus	63	7.2
2	227	–4.5	0.5	52.2	L	Supplementary motor area	126	6.6
					R	Supplementary motor area	101	6.6
3	42	46.5	–38.5	44.4	R	Inferior parietal lobule	26	6.4
					R	Supramarginal gyrus	16	6.5
4	16	13.5	–2.5	73	R	Supplementary motor area	10	6.4
					R	Superior frontal gyrus	6	6.8
5	14	36	–4	65.2	R	Superior frontal gyrus	13	6.7
6	11	4.5	8	54.8	R	Supplementary motor area	11	6.7
7	8	52.5	12.5	5.4	R	Inferior frontal gyrus, opercular part	8	6.2
8	51	–15	–19	8	L	Thalamus	51	7
9	28	–22.5	0.5	5.4	L	Putamen	26	6.5
10	24	–30	–7	–2.4	L	Putamen	23	6.5
11	22	–25.5	–23.5	5.4	L	Putamen	4	6.4
12	18	25.5	–2.5	5.4	R	Putamen	17	6.5
13	10	–19.5	5	21	L	Caudate nucleus	9	6.4
14	7	–24	–7	13.2	L	Putamen	6	6.4
15	7	–18	–7	23.6	L	Caudate nucleus	3	6.4
16	707	18	–46	–20.6	R	Hemispheric lobule IV/V	335	8
					R	Hemispheric lobule VI	161	7.6
					R	Vermic lobule IV/V	153	7.6
					R	Vermic lobule VI	34	6.6
					R	Vermic lobule VIII	3	6.5
17	288	–30	–56.5	–28.4	L	Hemispheric lobule VI	253	7.6
					L	Crus I	35	6.7
18	10	22.5	–65.5	–23.2	R	Hemispheric lobule VI	10	6.3
19	8	4.5	–61	–31	R	Vermic lobule VIII	6	6.5

We used two different paradigms to investigate the difference between the goal-directed and stimulus-driven motor acts. Figure 3 shows how distributions of activated neuronal clusters varied between hemispheres and large-scale parts of the human brain during these two types of movements. The stimulus-driven movements were characterized by strong leftward lateralization, whereas the goal-directed movements seem to involve activation of the bilateral cortex, basal ganglia, and cerebellum, with a pronounced rightward shift in the last case.

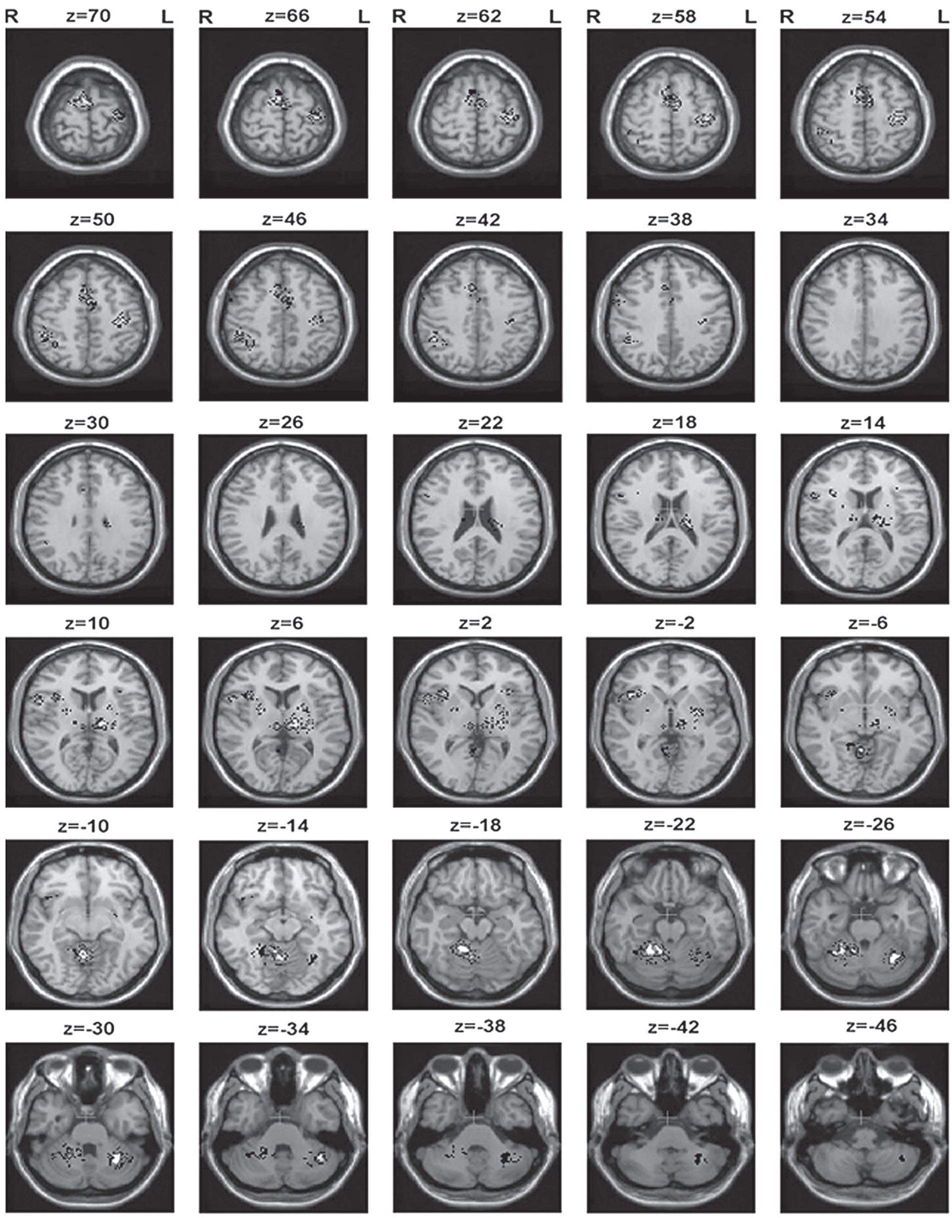


Figure 2. Areas of activation during goal-directed movement imposed on averaged brain

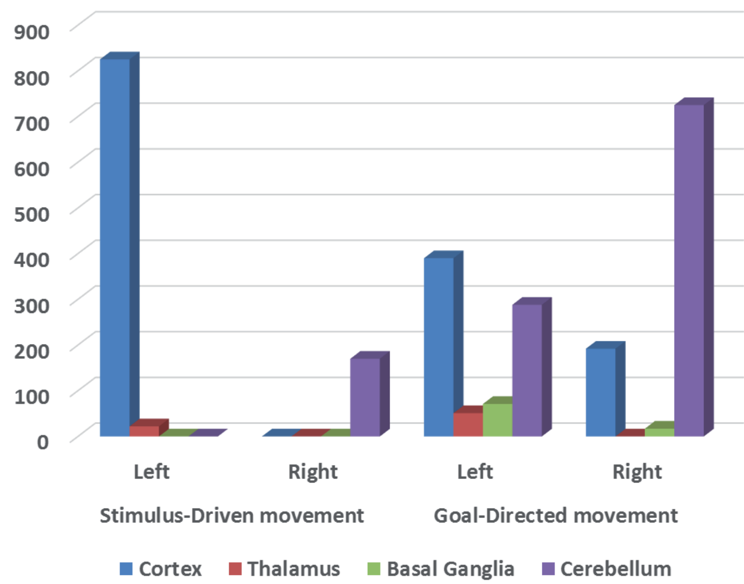


Figure 3. Comparison of brain activity (cluster size) during stimulus-driven and goal-directed movements

These findings suggest that goal-directed motor control is carried out by dispersed neural networks localized in both hemispheres. In particularly, cerebral activity seems to have the tendency to shift from motor to associative areas. Also worth noting is the participation of the bilateral posterolateral cerebellum in non-motor functions. This involvement significantly shifts rightward in the case of goal-directed movements.

Discussion

Analysis of human motor activity by means of event-related and block designed paradigms showed quite similar brain activation patterns in motor areas. As both conditions required subjects to perform movements, we can conclude that this common network, composed of cortical, subcortical, and cerebellar structures, is associated with motor function in general. Our finding is consistent with previous reports and suggests that the primary motor cortex (M1), which is the "lowest level" motor area for the control of motor acts, exerts influence, through pyramidal fibers passing down to the anterior horn's motor neurons, upon basal ganglia nuclei, which in turn exert extrapyramidal control of motor program sequences via the thalamus (Jueptner, & Weiller, 1998; Lanciego, Luquin, & Obeso, 2012). At the same time, the cerebellum is involved in equilibrium and the coordination and control of movement (Jueptner, & Weiller, 1998; Strick et al., 2009).

On the other hand, our study revealed significant differences in activity localization between the two motor paradigms, i.e., stimulus-driven and goal-directed. In the latter case, we observed activation of the bilateral supplementary motor area (SMA, pre-SMA), which is considered to play a role in the initiation of movement (Cunnington, Windischberger, Deecke, & Moser, 2003), and in action control

(Nachev, Wydell, O'Neill, Husain, & Kennard 2007). Furthermore, along with the motor cortex, the associative frontal and parietal areas were also engaged in goal-directed movement.

Another main difference is a pronounced right sided lateralization of brain activity in the associative parietal areas, frontal cortex, and basal ganglia during goal-directed motor activity. We suppose that this could be a manifestation of cognitive components of voluntary movement. Previous studies reviewed evidence for partially segregated networks of brain areas that carry out goal-directed and stimulus-driven attentional functions (Corbetta, & Shulman, 2002). One possible explanation might be that this lateralization is due to a close relationship between goal-directed motor behavior and voluntary attention. Specifically, the system which is thought to direct attention to behaviorally relevant stimuli is strongly lateralized to the right hemisphere (Shulman, Pope, Astafiev, McAvoy, Snyder, & Corbetta, 2010).

The most interesting fact is that goal-directed movement was associated with bilateral activation of the cerebellum along with the cortex. Bilateral cerebellar activation in lobules VI and Crus I during the n-back test was reported in recent studies, showing lateral cerebellar posterior lobe activation during working memory tests (Honey, Bullmore, & Sharma, 2000; Tomasi, Caparelli, Chang, & Ernst, 2005; Valera, Faraone, Biederman, Poldrack, & Seidman, 2005). Obviously, working memory capacity is an important feature of control and execution in attention-demanding tasks (Engle, Cantor, & Carullo, 1992; Kane, Bleckley, Conway, & Engle, 2001). Defined as the ability to maintain and manipulate information online in the absence of incoming sensory or motor stimulation, working memory can be one of the manifestations of internal model control (Ito, 2008). Recent evidence from neuroimaging and human lesion studies suggests that the right posterolateral cerebellar hemisphere is involved, independently of movement, in helping an individual to generate verbs for given nouns (Gebhart, Petersen, & Thach, 2002) and in the acquisition of a new lexicon (Lesage, Nailer, & Miall, 2016). The extent to which the cerebellar regions (right cerebellar vermis and right cerebellar Crus II), but not the cerebral areas, were recruited during learning correlated positively with participants' improvement in performance after the learning task. The data provide evidence for a cerebellar role not only in motor performance but in cognitive processing as well.

One of the reasons the cerebellum is involved in cognitive tasks is that movements themselves contain cognitive features. In our study, goal-directed movement before the signal requires internal timing, anticipation, and error correction. The activations during cognitive and emotional processing are localized to the cerebellar posterior lobe in lobules VI and VII, involving both Crus I and Crus II, with no anterior lobe involvement (Exner, Weniger, & Irle, 2004; Schmahmann, Weilburg, & Sherman, 2007; Tavano, Fabbro, & Borgatti, 2007). This suggests distinct, segregated cerebellar areas providing non-motor processing located in the posterior lobe. Activity in lobule VI was registered during a working memory task without any motor component (Stoodley, & Schmahmann, 2009). The existence of a significant lobule VI cluster in volunteers performing the simple clenching task in our study fits well with the idea that even nearly automatically produced movements preserve some residual cognitive properties.

Comparative anatomical studies show the enlargement of the ventral dentate and posterior cerebellar lobe in humans to be parallel to the enlargement of the prefrontal cortex (Leiner, Leiner, & Dow, 1991). These observations have led to the proposal that these areas must be related, and that posterolateral cerebellum participation in non-motor functions may be especially prominent in humans. Our neuroimaging data studies also prove that, as we see activation increase in frontal cortex areas like the SMA, the inferior frontal gyrus, and the opercular area, lobules VI and Crus I activate accordingly. This conjunction might reflect the shared function of these cerebral and cerebellar areas. The prefrontal cortex (PFC) receives input from all other cortical regions and functions to plan and direct motor, cognitive, affective, and social behavior. And as our activity (explicit and implicit) becomes more conditioned to social interaction and emotional state, the cerebellum, which was considered to be engaged solely in motor control, took on a wide range of non-motor functions, probably due to the development of new connections with the prefrontal and parietal areas (Takahashi et al., 2004).

Conclusion

In the present study, we have shown differences in the localization of the brain's movement-related activity, depending on the involvement of cognitive functions. These differences testify to the role of the right hemisphere and the cerebellum in motor cognition. In particular, our results suggest that right associative cortical areas together with the right posterolateral cerebellum (Crus I and lobule VI) and basal ganglia define cognitive control over motor activity, promoting the shift from stimulus-driven to goal-directed mode of processing. These results, along with recent data from research on cerebro-cerebellar circuitry, redefine the scope of future tasks for exploring the relatively unexpected contribution of the right hemisphere and especially the cerebellum to diverse aspects of human behavior and cognition.

Acknowledgments

This study was in part supported by the Russian Science Foundation (RScF 17-78-30029 — The brain architecture of semantic representations) and by the Russian Foundation for Basic Research (RFBR grant 15-04-05313, and ofi-m grants 15-29-01344, 17-29-02518 as related to brain mechanisms of attentional control and to the large-scale network organization, respectively).

References

Allen, G.I., & Tsukahara, N. (1974). Cerebro-cerebellar communication systems. *Physiological Reviews*, *54*(4), 957–1006.

Anand, B.K., Malhotra, C.L., Singh, B., & Dua, S. (1959). Cerebellar projections to limbic system. *Journal of Neurophysiology*, *22*(4), 451–457.

Appollonio, I. M., Grafman, J., Schwartz, V., Massaquoi, S., & Hallett, M. (1993). Memory in patients with cerebellar degeneration. *Neurology*, *43*(8), 1536–1536. doi: 10.1212/WNL.43.8.1536

Behroozmand, R., Shebek, R., Hansen, D.R., Oya, H., Robin, D.A., Howard, M.A., & Greenlee, J.D. (2015). Sensory–motor networks involved in speech production and motor control: An fMRI study. *Neuroimage*, *109*, 418–428. doi: 10.1016/j.neuroimage.2015.01.040

Botez-Marquard, T., Léveillé, J., & Botez, M. I. (1994). Neuropsychological functioning in unilateral cerebellar damage. *Canadian Journal of Neurological Sciences*, *21*(4), 353–357. doi: 10.1017/S0317167100040956

Corbetta, M., & Shulman, G.L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, *3*(3), 201–215. doi: 10.1038/nrn755

Cunnington, R., Windischberger, C., Deecke, L., & Moser, E. (2003). The preparation and readiness for voluntary movement: A high-field event-related fMRI study of the Bereitschafts-BOLD response. *Neuroimage*, *20*(1), 404–412. doi: 10.1016/S1053-8119(03)00291-X

Eickhoff, S.B., Stephan, K.E., Mohlberg, H., Grefkes, C., Fink, G.R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage*, *25*(4), 1325–1335. doi: 10.1016/j.neuroimage.2004.12.034

Engle, R.W., Cantor, J., & Carullo, J.J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 972–992. doi: 10.1037/0278-7393.18.5.972

Exner, C., Weniger, G., & Irle, E. (2004). Cerebellar lesions in the PICA but not SCA territory impair cognition. *Neurology*, *63*(11), 2132–2135. doi: 10.1212/01.WNL.0000146197.44568.CD

Gebhart, A.L., Petersen, S.E., & Thach, W.T. (2002). Role of the posterolateral cerebellum in language. *Annals of the New York Academy of Sciences*, *978*, 318–333. doi: 10.1111/j.1749-6632.2002.tb07577.x

Gentsch, A., Weber, A., Synofzik, M., Vosgerau, G., & Schütz-Bosbach, S. (2016). Towards a common framework of grounded action cognition: Relating motor control, perception and cognition. *Cognition*, *146*, 81–89. doi: 10.1016/j.cognition.2015.09.010

Grefkes, C., Eickhoff, S.B., Nowak, D.A., Dafotakis, M., & Fink, G.R. (2008). Dynamic intra-and interhemispheric interactions during unilateral and bilateral hand movements assessed with fMRI and DCM. *Neuroimage*, *41*(4), 1382–1394. doi: 10.1016/j.neuroimage.2008.03.048

Honey, G.D., Bullmore, E.T., & Sharma, T. (2000). Prolonged reaction time to a verbal working memory task predicts increased power of posterior parietal cortical activation. *NeuroImage*, *12*(5), 495–503. doi: 10.1006/nimg.2000.0624

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, *9*(4), 304–313. doi: 10.1038/nrn2332

Jahanshahi, M., Obeso, I., Rothwell, J. C., & Obeso, J. A. (2015). A fronto-striato-subthalamic-pallidal network for goal-directed and habitual inhibition. *Nature Reviews Neuroscience*, *16*(12), 719-732. doi: 10.1038/nrn4038

Jueptner, M., & Weiller, C. (1998). A review of differences between basal ganglia and cerebellar control of movements as revealed by functional imaging studies. *Brain*, *121*(8), 1437–1449. doi: 10.1093/brain/121.8.1437

Kane, M.J., Bleckley, M.K., Conway, A.R., & Engle, R.W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General*, *130*(2), 169. doi: 10.1037/0096-3445.130.2.169

Lanciego, J.L., Luquin, N., & Obeso, J.A. (2012). Functional neuroanatomy of the basal ganglia. *Cold Spring Harbor perspectives in medicine*, *2*(12), a009621. doi: 10.1101/cshperspect.a009621

Leiner, H.C., Leiner, A.L., & Dow, R.S. (1991). The human cerebro-cerebellar system: Its computing, cognitive, and language skills. *Behavioural Brain Research*, *44*(2), 113–128. doi: 10.1016/S0166-4328(05)80016-6

Lesage, E., Nailer, E.L., & Miall, R.C. (2016). Cerebellar BOLD signal during the acquisition of a new lexicon predicts its early consolidation. *Brain and Language*, *161*, 33–44. doi: 10.1016/j.bandl.2015.07.005

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., & Burdette, J.H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage*, *19*(3), 1233–1239. doi: 10.1016/S1053-8119(03)00169-1

Nachev, P., Wydell, H., O'Neill, K., Husain, M., & Kennard, C. (2007). The role of the pre-supplementary motor area in the control of action. *Neuroimage*, *36*, T155–T163. doi: 10.1016/j.neuroimage.2007.03.034

Nelson, S. M., Dosenbach, N. U., Cohen, A. L., Wheeler, M. E., Schlaggar, B. L., & Petersen, S. E. (2010). Role of the anterior insula in task-level control and focal attention. *Brain Structure and Function*, *214*(5–6), 669–680. doi: 10.1007/s00429-010-0260-2

Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-Oroz, M.C., Lehericy, S., Bergman, H. & Obeso, J.A. (2010). Goal-directed and habitual control in the basal ganglia: Implications for Parkinson's disease. *Nature Reviews Neuroscience*, *11*(11), 760–772. doi: 10.1038/nrn2915

Schmahmann, J.D., & Sherman, J.C. (1998). The cerebellar cognitive affective syndrome. *Brain*, *121*(4), 561–579. doi: 10.1093/brain/121.4.561

Schmahmann, J.D., Weilburg, J.B., & Sherman, J.C. (2007). The neuropsychiatry of the cerebellum — insights from the clinic. *Cerebellum*, *6*(3), 254–267.

Shulman, G.L., Pope, D.L., Astafiev, S.V., McAvoy, M.P., Snyder, A.Z., & Corbetta, M. (2010). Right hemisphere dominance during spatial selective attention and target detection occurs outside the dorsal frontoparietal network. *Journal of Neuroscience*, *30*(10), 3640–3651. doi: 10.1523/JNEUROSCI.4085-09.2010

Stoodley, C. J., & Schmahmann, J. D. (2009). Functional topography in the human cerebellum: A meta-analysis of neuroimaging studies. *Neuroimage*, *44*(2), 489-501. doi: 10.1016/j.neuroimage.2008.08.039

Strick, P.L., Dum, R.P., & Fiez, J.A. (2009). Cerebellum and nonmotor function. *Annual Review of Neuroscience*, *32*, 413–434. doi: 10.1146/annurev.neuro.31.060407.125606

Takahashi, H., Koeda, M., Oda, K., Matsuda, T., Matsushima, E., Matsuura, M., & Okubo, Y. (2004). An fMRI study of differential neural response to affective pictures in schizophrenia. *Neuroimage*, *22*(3), 1247–1254. doi: 10.1016/j.neuroimage.2004.03.028

Tavano, A., Fabbro, F., & Borgatti, R. (2007). Speaking without the cerebellum. In A. Schalley & D. Klentzos (Eds.). *Mental states: Evolution, function, nature* (pp. 171–190). Amsterdam: John Benjamins Publishing Company. doi: 10.1075/slcs.92.11tav

Tomasi, D., Caparelli, E.C., Chang, L., & Ernst, T. (2005). fMRI-acoustic noise alters brain activation during working memory tasks. *Neuroimage*, *27*(2), 377–386. doi: 10.1016/j.neuroimage.2005.04.010

Valera, E.M., Faraone, S.V., Biederman, J., Poldrack, R.A., & Seidman, L.J. (2005). Functional neuroanatomy of working memory in adults with attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *57*(5), 439–447. doi: 10.1016/j.biopsych.2004.11.034

Von der Gablentz, J., Tempelmann, C., Münte, T.F., & Heldmann, M. (2015). Performance monitoring and behavioral adaptation during task switching: An fMRI study. *Neuroscience*, *285*, 227–235. doi: 10.1016/j.neuroscience.2014.11.024

Wardman, D.L., Gandevia, S.C., & Colebatch, J.G. (2014). Cerebral, subcortical, and cerebellar activation evoked by selective stimulation of muscle and cutaneous afferents: An fMRI study. *Physiological Reports*, *2*(4), e00270.

Wu, T., Kansaku, K., & Hallett, M. (2004). How self-initiated memorized movements become automatic: A functional MRI study. *Journal of Neurophysiology*, *91*(4), 1690–1698. doi: 10.1152/jn.01052.2003

Original manuscript received August 25, 2017
Revised manuscript accepted September 11, 2017
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society | Lomonosov Moscow State University

Brain mechanisms of the Tip-of-the-Tongue state: An electroencephalography-based source localization study

Stanislav A. Kozlovskiy*, Sophie D. Shirenova, Anastasia K. Neklyudova, Alexander V. Vartanov

Faculty of Psychology, Lomonosov Moscow State University, Moscow, Russia

* Corresponding author. E-mail: stas@psy.msu.ru

Background. The Tip-Of-the-Tongue (TOT) state occurs when a person fails to retrieve a familiar word, e.g., a name, from long-term memory, while knowing perfectly well that the forgotten word exists in memory and being able to report some information about it (semantic associations, the first letter, the number of syllables, etc.).

Objective and Method. In the present work, we studied the activation of brain structures during the TOT state. The participants (N = 20; age 21.5 ± 4.1) viewed portraits of movie stars whose names they were asked to remember. Event related potentials (ERP) were registered in three conditions: 1) the participant remembered the name; 2) the participant did not know the name; 3) the participant knew the name but could not remember it (TOT-state). The sources of cortical activation were computed (dSPM algorithm).

Results. Time intervals demonstrating significant differences (t-test) in activation among the three conditions were calculated for each activated area, so that up to four different stages of processing could be delineated. According to our analysis, face perception involves activation of the visual cortex (left cuneus and right precuneus cortices), banks of the superior temporal sulci, poles of frontal and temporal lobes, and fusiform gyrus. The early activation does not depend on the successful retrieval of the name. A second increase in activation of the visual cortex is present at a later stage of processing, when name retrieval fails or if it is impeded.

Conclusion. We have shown that successful face recognition involves activation of the posterior cingulate cortex and the isthmus of the cingulate cortex in both hemispheres. Additionally, the parahippocampal gyrus is less active at the early stages and more active at the later stages of processing in the TOT-state, when name retrieval from the long-term memory fails.

Keywords: tip-of-the-tongue (TOT), memory retrieval, verbatim recollection, electroencephalography (EEG), source localization, event related potentials (ERP), posterior cingulate cortex, parahippocampal gyrus, isthmus of cingulate gyrus

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0315
http://psychologyinrussia.com

Introduction

When a person fails to remember a familiar word, a specific subjective state may appear which is known as the Tip-Of-the-Tongue (TOT) phenomenon. In this state, the person is able to retrieve some information about the forgotten word (its first letter, the number of syllables, etc.). S/he can also recall synonyms, which are are immediately rejected as incorrect (Freud, 1904) and correctly report some morphologic characteristics of the word, such as the position of the accented syllable (Brown & McNeill, 1966). This phenomenon is non-specific to culture and can be observed worldwide. Its frequency of occurrence increases with age: elderly people experience it more often (Salthouse & Mandell, 2013).

Although this phenomenon has been known for a long time, the first person who shifted the focus of attention to it and popularized such a feature of human memory was the Russian writer Anton Chekhov, who mentioned TOT in his humorous novel "A Horsey Name" (Chekhov, 1885). The main character could not recall a surname, but remembered that it was somehow related to horses. Then he tried to remember every surname that seemed associated with horses; in the end the answer turns out to be "Hayes". In the scientific literature, the first author who described the TOT phenomenon was William James (James, 1890). There are a variety of models in contemporary cognitive science that try to explain this phenomenon. Up to five different theories are currently known, which can be roughly divided into two groups (Schwartz, 1999).

The first three theories can be categorized as "direct-access approaches". They are focused on the problem of memory access. *The blocking hypothesis* (Jones, 1989) suggests that TOT occurs because a wrong word which sounds very similar comes to mind and blocks the retrieval of the correct word. *The incomplete activation theory* (Brown & McNeill, 1966) states that TOT occurs when the activation of the target word's phonological representation is poorer than the activation of words similar in meaning or in pronunciation. Although the target word is not retrieved, a subject can sense its presence in the mind. According to *the transmission deficit model* (Burke, MacKay, Worthley, & Wade, 1991), the verbatim recollection is a two-stage process. At the first stage, the semantic level of memory is activated where the core of the concept is recalled. Then the phonological level activates and the meaning of the word is translated into a verbal code. The TOT phenomenon occurs because of the failure of information transition from the semantic to the phonological levels.

The second group of theories emphasizes, instead of access to information in long-term memory interference with it. *The cue familiarity hypothesis* (Metcalfe, Schwartz, & Joaquim, 1993) suggests that the request to memory retrieval itself may seem to be very familiar such that a person has the feeling that the information is contained in his/her memory and so can be retrieved easily. According to this hypothesis, the TOT phenomenon emerges because of the difficulties of retrieval. Finally, *the accessibility heuristic theory* (Koriat, 1993; 1995) suggests that the more different information comes to mind while trying to recall the right word, the stronger the TOT-phenomenon is.

Although there are various experimental data and theoretical considerations in cognitive psychology related to this phenomenon, its brain mechanisms are not fully understood. Present psychophysiological studies mainly describe the characteristics of event-related potentials (ERPs) or brain rhythms. For example, it was shown that in the condition of a successful recall compared to the situation when a subject did not know the right name, differences are observed in P2, P3, and N450 of ERPs (Bujan, Galdo-Alvarez, Lindin, & Diaz, 2012). According to other authors (Resnik, Bradbury, Barnes, & Leff, 2014), positive depression of alpha-rhythm occurs during the TOT condition, which was interpreted as a manifestation of an ongoing search for semantic information. Psychophysiological data on activation of brain structures are still controversial. For instance, according to Galdo-Alvarez, Lindín, & Díaz (2011), during the retrieval of a name to match a face the following areas are active: the posterior temporal area, the insula, lateral and medial prefrontal areas, the medial temporal lobe, the anterior cingulate cortex (ACC), and the supplementary motor area (SMA). According to Lindín, Díaz, Capilla, Ortiz, and Maestú (2010), a slightly different list of activated regions was found during this process: left temporal and frontal areas, bilateral parahippocampal gyrus, right fusiform gyrus, bilateral occipital, left temporal, as well as right frontal and parietal areas.

The present study investigated the activation pattern of brain structures during the TOT phenomenon while recollecting a name to match a face. The data were compared to brain activation patterns in control conditions of successful name retrieval and perception of unfamiliar names.

Method

Twenty volunteers (age 21.5 ± 4.1; 10 males and 10 females) were recruited for this study. All participants were right-handed. They were healthy, with normal or corrected-to-normal vision acuity, with no history of neurological or psychiatric disorders. All participants gave their informed consent prior to their inclusion in the study, and did not report fatigue due to insufficient sleep. None of the participants were familiar with the protocols used in the study.

Seventy portraits of contemporary British and Hollywood movie stars (35 male and 35 female celebrities) were selected on the basis of normative familiarity judgments obtained by asking independent subjects. The participants viewed the stimuli from a distance of 70 cm. The angular size of the stimuli was 25 angular degrees in height and 16 angular degrees in width. All the stimuli were monochromatic. The background color was black. The actors were photographed face forward. All the pictures were adjusted by eye level and brought to a common format. The faces had no emotional expression (see Figure 1).

Figure 1. Examples of the actors' faces which were shown to the participants

The participants were asked to remember the names of the actors. They responded by pressing one of the three buttons:

1. Answer “Know”: *I remember the name*;
2. Answer “DK” [don’t know]: *I do not know the name*;
3. Answer “TOT” [tip-of-the-tongue]: *I knew the name, but I don’t remember it now*.

In order to make the responses automatic the participants were first asked to complete a training session. The task in this session was the same as in the main one. A different set of stimuli was presented for the training session: pictures of famous Russian and Soviet actors (15 male and 15 female faces). The training session lasted 4 minutes. The main session lasted 15 minutes.

EEG registration procedure. Each photograph appeared on the screen for 800 ms, followed by a 2,000 ms pause. For this time period, a fixation point was presented on a black screen. During the pause, the participant gave one of the three possible answers. The stimuli were presented in quasi-random order. Each photograph was shown four times (120 presentations in the training session and 280 presentations in the main session). During the sessions, the EEG registration was executed according to the 10-20% system using 19 active channels. The stimuli were shown via Presentation® software (version 18.2; Neurobehavioral Systems, Inc.; Berkeley, CA). The event-related potentials (ERPs) were averaged for the three conditions (“DK”, “Know”, “TOT”) for each participant.

Source localization procedure. The 3D-coordinates of the sources of brain activity were computed for the participants individually for each condition using the dSPM algorithm (Dale et al, 2000). Computed source coordinates were averaged and applied to an averaged brain surface anatomy model (ICBM152) using Brainstorm Software (Tadel, Baillet, Mosher, Pantazis, & Leahy, 2011; http://neuroimage.usc.edu/brainstorm). For each of the three possible answers, graphs were generated of activation dynamics for the 34 gyral-based regions of interest (ROIs) in each hemisphere. The outlines of the ROIs were determined according to the coordinates of the Desikan-Killiany labeling system (Desikan et al., 2006).

Only the ROIs that were highly active were selected for further analysis (12 brain areas in each hemisphere, see Figure 2): the banks of the superior temporal sulcus, the cuneus cortex, the entorhinal cortex, the frontal pole, the fusiform gyrus, the isthmus of the cingulate cortex, the lingual gyrus, the parahippocampal

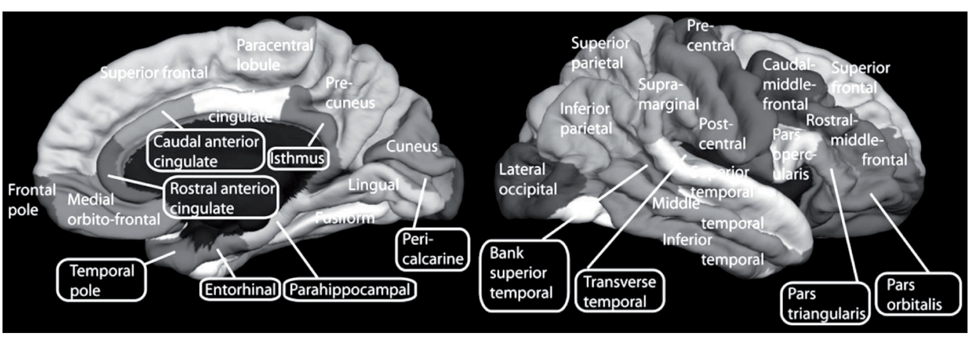


Figure 2. The cortex areas relevant for this study (after graphical materials of Hagmann et al., 2008; with permission of *Creative Commons Attribution License*)

gyrus, the pericalcarine cortex, the posterior cingulate cortex, the precuneus cortex, and the temporal pole.

Statistical analysis. For each ROI, time periods of significant differences in activation among three conditions were calculated using Student T-test for independent samples ($p < 0.05$; $N = 20$).

Results

No significant differences were found in either hemisphere for the banks of the superior temporal sulci, the poles of frontal and temporal lobes, and the fusiform gyri. T-test also did not reveal any significant differences between the conditions in activation of the left cuneus cortex and the right precuneus cortex (see Figure 3).

It can be observed that at early stages, the brain areas associated with primary visual preprocessing (the cuneus and the precuneus cortices) were active. The fusiform gyrus was active for around 150 ms after the stimulus. There was high activation of the temporal lobes (the temporal poles and the banks of the superior temporal sulci). Notably, in the left hemisphere the activation of the temporal poles was greater than in the right hemisphere. The activation of the frontal lobes increased during the whole time period that was analyzed.

Table 1. Time periods of significantly differing activation among the conditions (Student T-test). L – left hemisphere; R – right hemisphere

Cortex area	Hemi-sphere	TOT & DK (t, ms)	TOT & Know (t, ms)	DK & Know (t, ms)
Cuneus cortex	R	–	477–523	501–525, 558–568
Precuneus cortex	L	527–542, 688–696	487–568, 614–640	506–517, 552–565, 692–698, 734–744
Pericalcarine cortex	L	–	–	467–480
	R	468–476, 645–655, 761–767	–	464–481
Lingual gyrus	L	469–478,760–766	–	463–487
	R	464–477, 488–499	–	471–480
Parahippo-campal gyrus	L	362–367, 458–477	304–313	–
	R	336–363	–	–
Entorhinal cortex	L	–	303–314	–
	R	322–327	–	–
Isthmus of the cingulate cortex	L	459–465	–	507–519, 649–663, 693–701
	R	–	–	540–549
Posterior cin-gulate cortex	L	452–465	–	649–656
	R	452–462	–	641–658

Time periods (in ms) during which significant differences among the conditions ("DK", "TOT", "Know") were found are given in Table 1.

Significant differences among the three conditions were found in activation of the pericalcarine cortex, the lingual gyrus, the parahippocampal gyrus, the entorhinal cortex, the isthmus of the cingulate cortex, the posterior cingulate cortex in both hemispheres, and also the right cuneus cortex and the left precuneus cortex (see Figure 4).

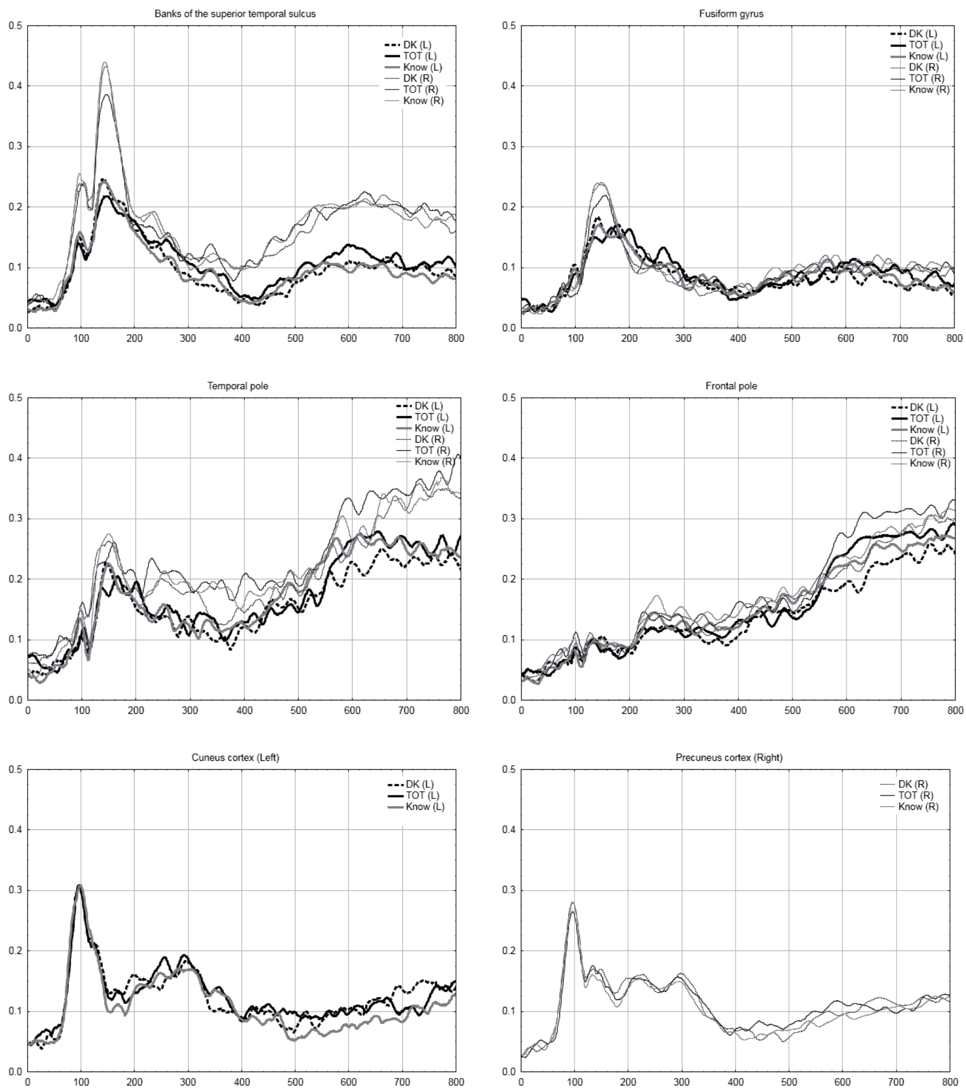


Figure 3. Activation of the brain areas for which no significant differences were found among the conditions. The data was averaged for the participants. The black line — "TOT", the dashed line — "DK", the gray line — "Know". Thick lines — left hemisphere, thin lines — right hemisphere

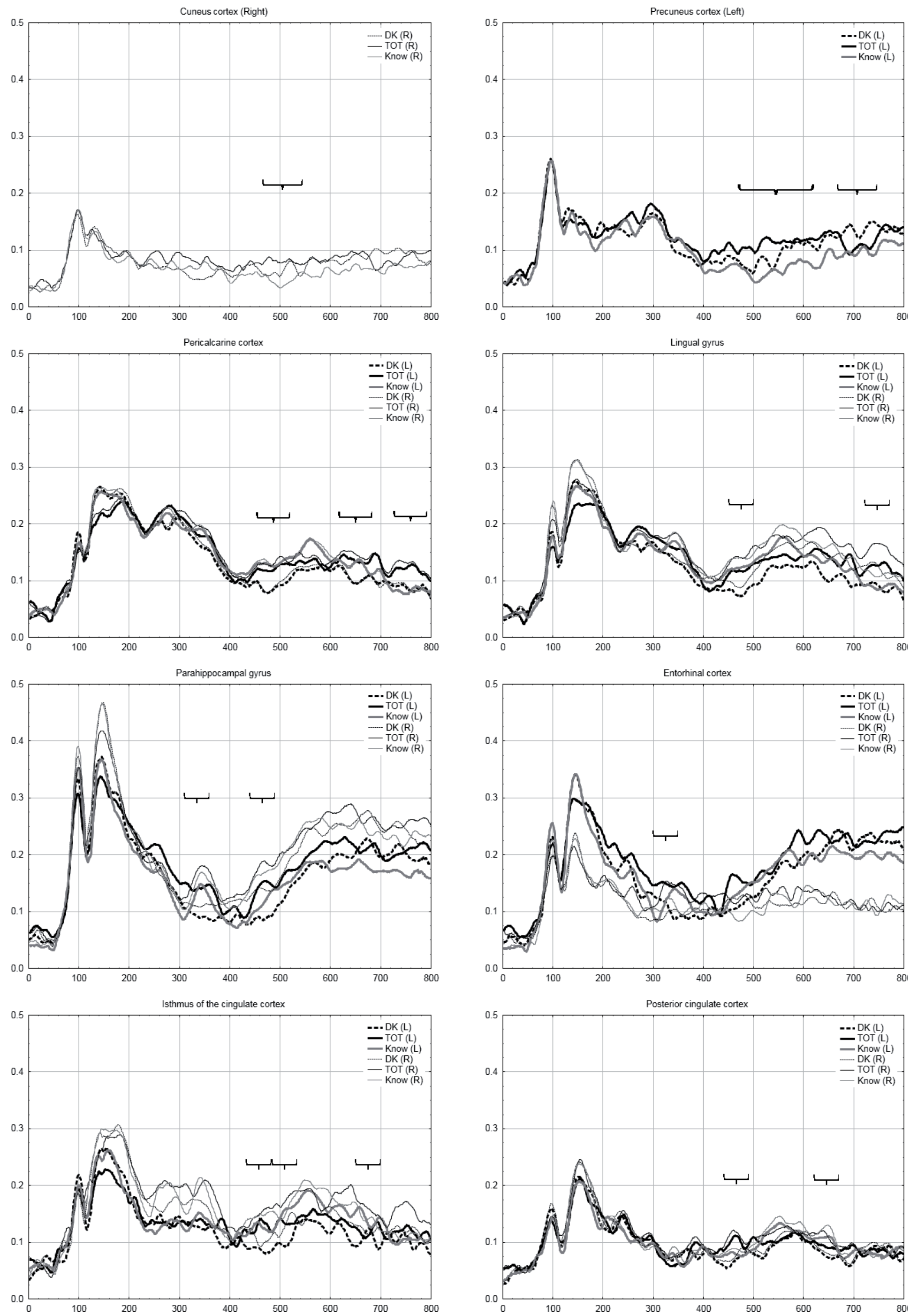


Figure 4. Activation of the brain areas for which significant differences were found among the conditions. The data was averaged for the participants. The black line — "TOT", the dashed line — "DK", the gray line — "Know". Thick lines — left hemisphere, thin lines — right hemisphere. The "└┬┘" sign marks time periods of significant differences among the conditions.

Discussion

It can be seen that during the analyzed time period (800 ms from the stimulus), activation of the brain structures occurs repeatedly and is relatively synchronized in different brain regions. According to our data, there are four relatively distinct stages when bursts of activation are present: around 100 ms, around 150-200 ms, around 300 ms, and around 400-500 ms. At the first and second stages, no significant differences (t-test) were found in activation of the brain structures among the conditions. The only exception, which does not follow the "four stages" model, is the activation of the frontal poles. It arises almost linearly during the time period analyzed (although it also has faint peaks at the latencies mentioned above). Together with the previously obtained data (Koechlin, 2011), this finding suggests that the activation of the frontal pole should be associated with the fact that the participants had to hold the instruction in their memory and control the task performance. The stages are discussed below.

The first stage is a steep and short increase in activation of the brain structures around 100 ms from the appearance of the stimulus. This peak is well defined in the cuneus and the precuneus cortices, where it is the highest peak of the analyzed time period. Notably, in the precuneus cortex there were no significant differences between the left and right hemisphere, while the cuneus cortex in the left hemisphere is more active than in the right hemisphere. We suggest that such early activation of the medial occipital and medial inferior parietal cortex at this stage reflects signal detection and its preprocessing. Activation is also present in several other brain regions at this stage (e.g., the lingual gyrus, the parahippocampal gyrus, the entorhinal cortex, the posterior cingulate cortex). However, in these structures this peak is significantly fainter than at the next stages. This may provide evidence that the appearance of the stimulus itself alerts these brain regions, preparing them for the following processing of the stimulus.

The second stage occurs at around 150-200 ms after the stimulus presentation. It is described by a steep increase in activation, which reaches its maximum at 150 ms; then the activation slowly decreases by 200 ms. This is the stage of the maximum activation of the brain. Significant differences (t-test) among the conditions are also absent at this stage. This peak is best defined in the banks of the superior temporal cortex and the parahippocampal gyrus. It is also well defined in the pericalcarine cortex, the lingual gyrus, the fusiform gyrus, the entorhinal cortex, the isthmus of the cingulate cortex, and the posterior cingulate cortex.

Notably, the fusiform gyrus is active only at this stage. During the rest of the analyzed time period, the activation of this brain region does not exceed the noise level. It was shown in multiple studies that the fusiform gyrus is involved in face perception (for a review, see Weiner & Zilles, 2016). This corresponds with our results, since we used faces as stimuli.

Activation of the banks of the superior temporal sulcus at this stage is higher in the right hemisphere (it is also present in the left hemisphere, but there it is almost half as low in value). While this region has been fully investigated in the left hemisphere (Wernike's area), its function in the right hemisphere requires further analysis. Previous data suggests that this brain region is active, together with the medial orbitofrontal cortex, during a face attractiveness assessment task (O'Doherty, Winston, Critchley, Perrett, Burt, & Dolan, 2003; Kranz & Ishai, 2006). We did not observe such a coordinated activation in our experiment. A further hypothesis (Karnakh, 2001) can explain the results. It states that the right superior temporal sulcus is involved in the interfacing of the dorsal and ventral streams of visual processing. This suggestion was based on a study of clinical cases of visual spatial neglect. Accordingly, activation of the banks of the superior temporal sulcus in our experiment could be explained as a manifestation of visual stimulus processing. A recent elaboration of this view relates the region's activity to the hippocampal formation (Velichkovsky, Krotkova, Sharaev, & Ushakov, 2017). Indeed, the parahippocampal gyrus is highly active at this stage, more so than any other analyzed brain region during the whole experiment.

In the well-known study of Quiroga et al. (Quiroga, Reddy, Kreiman, Koch, & Fried, 2005), where recording electrodes were implanted into the hippocampus, the parahippocampal gyrus, and the entorhinal cortex, it was found that neurons in this area are selectively activated by the faces of specific people. We used photos of famous actors, so the participants had viewed them multiple times regardless of whether they knew their names. Thus, in our experiment the activation of the parahippocampal gyrus may be caused by perception of a familiar face and activation of the information about this actor in long-term memory. Notably, the activation of the parahippocampal gyrus in the TOT state in both hemispheres at this stage was lower than in both other conditions. This finding can testify to an incomplete activation of the semantic network while perceiving a familiar face, with resulting inability to retrieve a name. This interpretation conforms to "the incomplete activation theory" (Brown & McNeill, 1966), according to which the TOT phenomenon is caused by a low activation of memory traces. However, at the later stages, the activation of this brain region is higher for the TOT condition, which contradicts this theory.

At the third stage a local maximum appears at around 300 ms from the stimulus, which slowly decreases by 400 ms. This peak is lower in amplitude than the first and second peaks described above. It was observed in activation of the visual cortex structures (the left cuneus cortex, the precuneus cortex, the pericalcarine cortex, and the lingual gyrus). Moreover, in the "DK" condition, this activation is the lowest, and in the "TOT" condition the activation is the highest. This secondary activation of the occipital region could be explained as visual cortex re-entrance (Ivanitskiy, 1976; Edelman, 1989). Apparently, subjects at this stage examine a face with more focal attention, making sure that they have perceived it correctly. This peak is also well expressed in the parahippocampal gyrus and the entorhinal cortex, although in this case, the activation significantly differs (t-test), depending on whether the name retrieval was successful or not. In the "DK" condition, when a participant decided that s/he doesn't know the target name, this peak is not present in either of these two structures. On the contrary, in the other two conditions this peak of activation was clearly observed. One possible explanation is that at the previous stage a participant who saw and recognized a familiar face, decided that s/he does not know the target name, and that there is no need to keep searching for it in long-term memory. The activation observed in the "Know" and in the "TOT" conditions may reflect a search for a name of a recognized person.

The fourth stage is the last burst of activation of the brain regions during the analyzed time period, and the one of the longest duration. The increase in activations begins around 400 ms from the stimulus, reaches its maximum by 600 ms from the stimulus, and slowly decreases by 800 ms. It consists of a 6-to-8 range of faint peaks of activation. At this stage a secondary activation of the visual cortex was observed (in the lingual gyrus, the pericalcarine cortex, the precuneus cortex, and faintly in the cuneus cortex). Notably, the highest activation was observed in the TOT condition, when a participant failed to retrieve the target name. Apparently, in this case a participant failing to remember the name looks closely at the face again hoping to find a clue to the name in the face of the actor. This hypothesis explains the known fact of alpha-rhythm depression that occurs during the TOT state (Resnik, Bradbury, Barnes, & Leff, 2014). The parahippocampal gyrus, which was discussed above, is also active, whereby its activation is higher in the TOT condition. This may be explained as an additional attempt to remember the target name. Another explanation of this finding is provided by the *accessibility heuristic theory* (Koriat, 1993; 1995), described above.

In all three conditions, high activation of the temporal poles was observed, especially in the right hemisphere. According to a meta-analysis (Olson, Plotzker, & Ezzyat, 2007), this brain region is involved in face perception, but it is mostly associated with socio-emotional processing. In our experiment, activation of the temporal poles may reflect an emotional attitude towards the actor.

At this last stage, the activation of the isthmus of the cingulate cortex and the posterior cingulate cortex was observed, with significant differences among the three conditions. The lowest activation was observed in the "DK" condition and the highest in the "Know" condition. Apparently these brain regions detect whether the retrieval from the long-term memory was successful. This explanation corresponds with data of a previous study (Kozlovskiy, Vartanov, Nikonova, Pyasik, & Velichkovsky, 2012, 2013) demonstrating that the volume of these brain areas negatively correlates with the number of memory errors. It has also been shown that there are strong causal interactions between the posterior cingulate and the hippocampal formation (Ushakov et al., 2016).

Conclusion

Retrieving a name by a photograph of a person is a complex task, which consists of multiple interacting cognitive and affective processes, such as visual recognition of a stimulus as a face, assessment of its attractiveness and its emotional expression, recognition of the face, name search in the verbal memory, decision making, and cognitive control. According to our data, this task activates multiple brain regions associated with various brain functional systems: the banks of the superior temporal sulcus, the cuneus cortex, the entorhinal cortex, the frontal pole, the fusiform gyrus, the isthmus of the cingulate cortex, the lingual gyrus, the parahippocampal gyrus, the pericalcarine cortex, the posterior cingulate cortex, the precuneus cortex, and the temporal pole.

Face perception involves activation of the visual cortex (the left cuneus and the right precuneus cortices), the banks of the superior temporal sulci, the poles of frontal and temporal lobes, and the fusiform gyri, which does not depend on whether the retrieval of the name from long-term memory was successful. If name retrieval fails, a second increase in activation of the visual cortex is present at later time intervals (right cuneus cortex). We have shown that successful face recognition involves activation of the posterior cingulate cortex and isthmus of the cingulate cortex in both hemispheres. According to our results, the parahippocampal gyrus is less active at the early stages and more active at the later stages of processing, when name retrieval from long-term memory fails.

Unfortunately, the method that was used does not allow the analysis of processing within the subcortical structures. For instance, we were not able to analyze the activation of the hippocampus, the brain structure that is massively involved in memory task solution. Therefore, we plan to overcome these limitations by completing our experimental framework with data of a functional MRI study and an EEG study with a large number of active channels.

Acknowledgments

This study was partially supported by the Russian Science Foundation (project RScF16-18-00066). We wish to express our gratitude for the criticism and enriching discussion of our results by B.M. Velichkovsky.

References

Brown, R., & McNeil, D. (1966). The “tip of the tongue” phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5*(4), 325–337. doi: 10.1016/S0022-5371(66)80040-3

Bujan, A., Galdo-Alvarez, S., Lindin, M., & Diaz, F. (2012). An event-related potentials study of face naming: Evidence of phonological retrieval deficit in the tip-of-the-tongue state. *Psychophysiology*, *49*(7), 980–990. doi: 10.1111/j.1469-8986.2012.01374.x

Burke, D.M., MacKay, D.G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of memory and language*, *30*(5), 542–579. doi: 10.1016/0749-596X(91)90026-G

Chekhov, A.P. (1915). A horsey name. In M. Fell (Ed.), *Russian silhouettes; more stories of Russian life*. New York: C. Scribner’s sons. (Original work published 1885)

Dale, A.M., Liu A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., & Halgren, E. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron, 26*(1), 55–67. doi: 10.1016/S0896-6273-(00)81138-1

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D. … Killiany, R.J., (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, *31(3),* 968–980. doi: 10.1016/j.neuroimage.2006.01.021

Edelman, G.M. (1989). *The remembered present. A biological theory of consciousness*. New York: Basic Books.

Freud, S. (1904). *Zur Psychopathologie des Alltagslebens* [On psychopathology of everyday life]. Berlin: S. Karger.

Galdo-Alvarez, S., Lindín M., & Díaz F. (2011). Brain dynamics associated with face-naming and the tip-of-the-tongue state. *Psicothema, 23*(2), 189–195.

Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., & Sporns, O. (2008). Mapping the structural core of human cerebral cortex. *PLOS Biology, 6*(7), 1479–1493. doi: 10.1371/journal.pbio.0060159

Karnath, H.-O. (2001). New insights into the functions of the superior temporal cortex. *Nature Reviews Neuroscience, 2*(8), 568–576. doi: 10.1038/35086057

Kranz, F., & Ishai, A. (2006). Face perception is modulated by sexual preference. *Current Biology*, *16*(1), 63–68. doi: 10.1016/j.cub.2005.10.070

Koechlin, E. (2011). Frontal pole function: What is specifically human? *Trends in Cognitive Sciences*, *15*(6), 241. doi: 10.1016/j.tics.2011.04.005

Ivanitskiy, A.M. (1976). *Mozgovye mekhanizmy otsenki signalov* [Brain mechanisms of signal evaluation]. Moscow: Meditsina.

James, W. (1890). *The principles of psychology*, New York: H. Holt and Company. doi: 10.1037/11059-000

Jones, G.V. (1989). Back to Woodworth: Role of interlopers in the tip-of-the-tongue phenomenon. *Memory & Cognition, 17*(1), 69–76. doi: 10.3758/BF03199558

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100,* 609–639. doi: 10.1037/0033-295X.100.4.609

Koriat, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General, 124,* 311–333. doi: 10.1037/0096-3445.124.3.311

Kozlovskiy S.A., Vartanov A.V., Nikonova E.Y., Pyasik M.M., & Velichkovsky B.M. (2012). The cingulate cortex and human memory processes. *Psychology in Russia: State of the Art, 5,* 231–243. doi: 10.11621/pir.2012.0014

Kozlovskiy S.A., Vartanov A.V., Nikonova E.Y., Pyasik M.M., & Velichkovsky B.M. (2013). Anatomical characteristics of cingulate cortex and neuropsychological memory tests performance. *Procedia - Social and Behavioral Sciences, 86*(10), 128–133. doi: 10.1016/j.sbspro.2013.08.537

Lindín, M., Díaz, F., Capilla, A., Ortiz, T., & Maestú, F., (2010). On the characterization of the spatio-temporal profiles of brain activity associated with face naming and the tip-of-the-tongue state: A magnetoencephalographic (MEG) study. *Neuropsychologia, 48*(6), 1757–1766. doi: 10.1016/j.neuropsychologia.2010.02.025

Metcalfe, J., Schwartz, B.L., & Joaquim, S.G. (1993). The cue familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 19,* 851–861. doi: 10.1037//0278-7393.19.4.851

O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D.M., & Dolan, R.J. (2003). Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia, 41*(2), 147–155. doi: 10.1016/S0028-3932(02)00145-8

Olson, I.R., Plotzker, A., & Ezzyat, Y. (2007). The enigmatic temporal pole: A review of findings on social and emotional processing. *Brain*, *130*(7), 1718-1731. doi: 10.1093/brain/awm052

Resnik, K., Bradbury, D., Barnes, G.R., & Leff, A.P. (2014). Between thought and expression, a magnetoencephalography study of the "tip-of-the-tongue" phenomenon. *Journal of Cognitive Neuroscience, 26*(10), 2210–2223. doi: 10.1162/jocn_a_00611

Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., & Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature, 435*(7045), 1102–1107. doi: 10.1038/nature03687

Salthouse, T.A., & Mandell A.R. (2013). Do age-related increases in tip-of-the-tongue experiences signify episodic memory impairments? *Psychological Science, 24*(12), 2489–2497. doi: 10.1177/0956797613495881

Schwartz, B.L. (1999). Sparkling at the end of the tongue: The etiology of tip-of-the-tongue phenomenology. *Psychonomic Bulletin & Review*, *6*(3), 379–393. doi: 10.3758/BF03196771

Sugiura, M., Shah, N.J., Zilles, K., & Fink, G.R. (2005). Cortical representations of personally familiar objects and places: Functional organization of the human posterior cingulate cortex. *Journal of Cognitive Neuroscience, 17*(2), 183–198. doi: 10.1162/0898929053124956

Tadel F., Baillet S., Mosher J.C., Pantazis D., & Leahy R.M. (2011). BrainStorm: a user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience, 2011,* 1–13. doi: 10.1155/2011/879716

Ushakov, V.L., Sharaev, M. G., Kartashov, S.I., Zavyalova, V.V., Verkhlyutov, V.M., & Velichkovsky, B.M. (2016) Dynamic causal modeling of hippocampal links within the human default mode network: Lateralization and computational stability of effective connections. *Frontiers in Human Neuroscience, 10,* 1–14. doi: 10.3389/fnhum.2016.00528

Velichkovsky, B.M., Krotkova, O.A., Sharaev, M.G., & Ushakov, V.L. (2017). In search of the "I": Neuropsychology of lateralized thinking meets Dynamic Causal Modeling. *Psychology in Russia: State of the Art.* This issue.

Weiner, K.S., & Zilles, K. (2016). The anatomical and functional specialization of the fusiform gyrus. *Neuropsychologia*, *83*(3), 48–62. doi: 10.1016/j.neuropsychologia.2015.06.033

Original manuscript received December 06, 2016
Revised manuscript accepted February 18, 2017
First published online September 30, 2017

Psychology in Russia: State of the Art
Volume 10, Issue 3, 2017

Russian Psychological Society
Lomonosov Moscow State University

Exploring *terra incognita* of cognitive science: Lateralization of gene expression at the frontal pole of the human brain

Irina A. Dolina[a], Olga I. Efimova[a,b], Evgeniy M. Kildyushov[c], Aleksey S. Sokolov[d], Philipp E. Khaitovich[b,e,f], ArtemV. Nedoluzhko[a], Fyodor S. Sharko[a], Boris M. Velichkovsky[a,h,i*]

[a] *National Research Center "Kurchatov Institute", Moscow, Russia*
[b] *Skolkovo Institute for Science and Technology, Skolkovo, Russia*
[c] *Pirogov Russian National Research Medical University, Moscow, Russia*
[d] *Limited Liability Company "Elgene", Krasnogorsk, Russia*
[e] *CAS-MPG Partner Institute for Computational Biology, Shanghai, China*
[f] *Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany*
[h] *Moscow Institute for Physics and Technology, Moscow, Russia*
[i] *Russian State University for the Humanities, Moscow, Russia*

* Corresponding author. E-mail: velich@applied-cognition.org

Background. Rostral prefrontal cortex, or frontopolar cortex (FPC), also known as Brodmann area 10 (BA10), is the most anterior part of the human brain. It is one of the largest cytoarchitectonic areas of the human brain that has significantly increased its volume during evolution. Anatomically the left (BA10L) and right (BA10R) parts of FPC show slight asymmetries and they may have distinctive cognitive functions. *Objective.* In the present study, we investigated differential expression of the transcriptome in the left and right parts of BA10.

Design. Postmortem samples of human brain tissue from fourteen donors (male/female without history of psychiatric and neurological diseases, mean age 39.79±3.23 years old, mean postmortem interval 12.10±1.76 h) were obtained using the resources of three institutions: the Partner Institute of Computational Biology of Chinese Academy of Sciences, the Max Planck Institute for Evolutionary Anthropology, and NIH Neuro-BioBank.

Results. By using a standard RNA-sequencing followed by bioinformatic analysis, we identified 61 genes with differential expression in the left and right FPC. In general, gene expression was increased in BA10R relative to BA10L: 40 vs. 21 genes, respectively. According to gene ontology analysis, the majority of up-regulated genes in BA10R belonged to the protein-coding category, whereas protein-coding and non-coding genes

ISSN 2074-6857 (Print) / ISSN 2307-2202 (Online)

doi: 10.11621/pir.2017.0316
http://psychologyinrussia.com

were equally up-expressed in BA10L. Most of the up-regulated genes in BA10R were involved in brain plasticity and activity-dependent mechanisms also known for their role in the hippocampus. 24 out of 30 mental disorder-related genes in the dataset were disrupted in schizophrenia. No such a wide association with other mental disorders was found.

Conclusion. Discovered differences point at possible causes of hemispheric asymmetries in the human frontal lobes and at the molecular base of higher-order cognitive processes in health and disease.

Keywords: neuropsychology, frontopolar cortex, human cerebral asymmetry, Yakovlevian torque, RNA transcriptome, sequencing, schizophrenia, attention

Introduction

Rostral prefrontal cortex, or frontopolar cortex (FPC), also known as Brodmann area 10 (BA10), is the most anterior part of the human brain. This area extended its complexity during hominid evolution by e.g. considerable increase in neuron numbers, a specific increase in connectivity and a dramatic increase in size: gibbon — 0.2 cm^3 (0.2%), bonobo — 2.8 cm^3 (0.7%), human — 14.2 cm^3 (1.2% of the whole brain volume) (Semendeferi, Armstrong, Schleicher, Zilles, & Van Hoesen, 2001). Functional neuroimaging studies of prefrontal cortex showed strong hemodynamic changes in this area under different conditions, from the simplest to the most complex tasks, such as language understanding or contemplating future actions, in accordance with internal goals and plans (Burgess, Dumontheil, & Gilbert, 2007; Miller, & Cohen 2001). Together with clinical observations, it appears that rostral prefrontal cortex, considered as a whole, supports higher cognitive processes such as planning, decision-making, retrieval of memories, establishing logical links and multi-tasking that allows performance of mental and physical activities at the same time or in close temporal succession (Braver, & Bongiolatti, 2002; Penfield, & Evans 1935; Ramnani, & Owen, 2004). In general, the left and right hemispheres of the human brain are anatomically asymmetric (LeMay, 1999) and probably support different cognitive functions. A number of studies suggests a particular role of the right FPC in self-referential rather than other-referential encoding (Craik et al., 1999; Christoff, & Gabrieli, 2000; Soch et al., 2016) and in understanding of concealed, or indirect meaning as in metaphoric speech, humor, irony and sarcasm (see however Forgacs, Lukács, & Pléh, 2014). The functional differences are echoed by anatomical data. Although basic mechanisms of language perception and production, traditionally considered as the *differentia specifica* of *Homo sapiens sapiens*, are localized in the left hemisphere, the right frontopolar area (BA10R) has a larger volume and demonstrates more rapid growth, both in anthropogenesis and in early ontogenesis (Hrvoj-Mihic, Bienvenu, Stefanacci, Muotri, & Semendeferi, 2013). This can be related to the Yakovlevian Torque phenomenon, in which frontal structures anterior to the right Sylvian fissure are 'torqued forward' relative to their counterparts on the left. The left occipital lobe is also splayed across the midline and skews the interhemispheric fissure in a rightward direction. First described by P.I Yakovlev, this phenomenon has been supported by fragmented paleoneurological findings (Toga, & Thompson, 2003).

Although hemispheric lateralization is not a feature unique to humans (Halpern, Gunturkun, Hopkins, & Rogers, 2005), the most publications on this asymmetry is dedicated to the anatomy, physiology and clinical pathology of the human brain (Herbert et al., 2005; Pujol et al., 2002; Toga, & Thompson, 2003). For example, cognitive disorders and psychiatric diseases such as schizophrenia, autism and dyslexia are accompanied by disturbances in brain asymmetry (Carper, Treiber, DeJesus, & Muller, 2016; Renteria, 2012; Y. Sun, Chen, Collinson, Bezerianos, & Sim, 2017).

Previous research has shown that the human genome produces a consistent molecular architecture in the cortex, despite millions of genetic differences across individuals and races (Colantuoni et al., 2011). However, molecular mechanisms leading to development of brain asymmetry in the adult human brain remain, with a few exceptions, underexplored (Kang et al., 2011; Karlebach, & Francks, 2015). In particular, molecular bases of the functional differences between distinct brain regions are either unclear (Pletikos et al., 2014) and disputed (T. Sun, Collura, Ruvolo, & Walsh, 2006), or complicated for subsequent analysis (Hawrylycz et al., 2012). This is especially true with respect to the evolutionary new prefrontal cortex.

In the present study, we applied the common RNA-Seq technique together with Gene Ontology (GO) analysis to investigate differential expression of the transcriptome in the left and right parts of Brodmann area 10 (BA10L and BA10R), i.e. around the poles of the human brain which show a significant diversity in anatomical and neuropsychological features. Understanding the differences in the transcriptome patterns of these most rostral cortical regions may have important basic and clinical relevance.

Method

Material and methods

Human brain tissue

Postmortem samples of human brain tissue from FPC areas in the left and right hemisphere (BA10L and BA10R) from fourteen donors were obtained using the resources of three institutions: the Partner Institute of Computational Biology of Chinese Academy of Sciences, the Max Planck Institute for Evolutionary Anthropology, and NIH NeuroBioBank. This sampling consisted of seven BA10L samples and seven BA10R samples from male/female individuals without known history of psychiatric and neurological diseases. Mean age 39.79±3.23 years old, mean postmortem interval 12.10±1.76 h (mean±SD).

RNA extraction

Isolation of total RNA from brain tissue samples was carried out using the Trizol reagent (Thermo Fisher Scientific, USA) by the standard technique. The quantity of total RNA was determined using a Nanodrop spectrophotometer (Thermo Fisher Scientific, USA). RNA integrity number (RIN) was assessed by BioAnalyzer 2100 (Agilent Technologies, USA) and RNA 6000 Nano kit (Agilent Technologies, USA). The RIN ranged from 6.1 to 9.3 for all samples.

Library preparation and sequencing

14 cDNA libraries for sequencing were constructed using 10 μg of RNA per sample and the mRNA-Seq Sample Prep Kit (Illumina, USA) according the manufacturer's protocol. The final library met all quality metrics as defined by Illumina, and library quantization was performed on an Agilent 2100 Bioanalyzer with a High-Sensitivity DNA kit (Agilent Technologies, USA) prior to sequencing. DNA-libraries were sequenced using an Illumina HiSeq 1500 platform (Illumina, USA) with 150-bp paired-end reads.

Sequencing analysis

DNA-reads were mapped to the reference human genome (hg19) using the TopHat program. The differential gene expression between BA10L and BA10R were analyzed using the Cufflinks package from the Cufflinks program (Trapnell et al., 2012). RPKM analysis (reads per kilobase per million mapped reads) was used as the normalized value of the expression level (Mortazavi et al., 2008).

Gene Ontology (GO) analysis was performed using the DAVID 6.8 (Database for Annotation, Visualization and Integrated Discovery) (Huang da, Sherman, & Lempicki, 2009) and PANTHER classification system (Mi et al., 2017). Cell type specific expression in the human cerebral cortex for studied genes was checked in the Human Protein Atlas database (version 4.1) and Ensembl (version 54.36) (Ponten et al., 2009). To investigate whether revealed genes with differential expression were constitutive or inducible, we used the database consisting of 3804 human housekeeping genes (Eisenberg, & Levanon, 2013). Finally, human disease enrichment and inducibility analysis of revealed genes was performed up to 30.03.2017 using the PubMed database and Schizophrenia Gene Resource (SZGR) (Jia, Han, Zhao, Lu, & Zhao, 2017).

Results

The total number of raw reads generated for both BA10L and BA10R areas was from 18 to 30 million sequences per sample (NCBI Bioproject: PRJNA388140). At least 90% of reads were mapped to the reference genome (except one sample). The vast majority of identified genes did not differ in expression between BA10L and BA10R areas. But some genes we found demonstrated significant differential expression.

The comparative study of the gene expression in the right and left parts of FPC allowed us to identify 61 genes that showed lateralization: 40 genes were up-regulated in BA10R whereas only 21 genes were up-regulated in BA10L Therefore, gene expression generally increased in BA10R relative to BA10L. All statistically significant cases of gene differential expression between BA10L and BA10R are listed in Table 1. Interestingly, the non-coding RNA (i.e. expression suppressors) and protein-coding genes were equally up-regulated in BA10L (10 and 11 genes, respectively), whereas the expression of protein-coding genes strongly dominated in BA10R (5 and 35 genes, respectively) (the *chi*-squared statistic is 9.1589, $df = 3$, $p \leq 0.01$).

Table 1. Statistically significant differential gene expressions in the human FPC

Gene	Brodmann areas	log2(fold change)	*p–value*	*q–value*
HSPA7	BA10L vs. BA10R	3.8592	0.00005	0.01981
RGS1	BA10L vs. BA10R	2.66997	0.0001	0.037021
IPO9	BA10L vs. BA10R	−2.10138	0.00005	0.01981
MIR34A+MIR34AHG	BA10L vs. BA10R	−2.25978	0.00005	0.01981
GAS5+SNORD80+SNORA103	BA10L vs. BA10R	−1.31268	0.00005	0.01981
BAG3	BA10L vs. BA10R	1.63766	0.00005	0.01981
ADM	BA10L vs. BA10R	1.89858	0.00005	0.01981
CD44	BA10L vs. BA10R	1.48236	0.00005	0.01981
NPAS4	BA10L vs. BA10R	2.51096	0.00005	0.01981
IFITM3	BA10L vs. BA10R	1.36025	0.00005	0.01981
H19+MIR675	BA10L vs. BA10R	−6.43921	0.00005	0.01981
USP15	BA10L vs. BA10R	−2.95303	0.00005	0.01981
MIR331+MIR3685	BA10L vs. BA10R	4.65469	0.00005	0.01981
TMCC3	BA10L vs. BA10R	1.98614	0.00005	0.01981
FOS	BA10L vs. BA10R	1.65584	0.00005	0.01981
SERPINA3	BA10L vs. BA10R	5.68007	0.00005	0.01981
MEG8+MIR370	BA10L vs. BA10R	−2.52033	0.00005	0.01981
HAUS4	BA10L vs. BA10R	−2.57307	0.00005	0.01981
MIR9-3HG	BA10L vs. BA10R	−1.10873	0.00005	0.01981
NFATC2IP	BA10L vs. BA10R	1.38301	0.00005	0.01981
MT1X	BA10L vs. BA10R	1.27587	0.0001	0.037021
Non-protein coding gene or gene fragment, no protein prediction available	BA10L vs. BA10R	inf	0.00005	0.01981
TAOK1	BA10L vs. BA10R	−2.56172	0.00005	0.01981
NSRP1	BA10L vs. BA10R	1.01791	0.00005	0.01981
ITGB4	BA10L vs. BA10R	1.93095	0.00005	0.01981
SLC14A1	BA10L vs. BA10R	1.34783	0.0001	0.037021
AMH+SF3A2	BA10L vs. BA10R	1.9429	0.00005	0.01981
CREB3L3	BA10L vs. BA10R	4.66498	0.00005	0.01981
CCDC61	BA10L vs. BA10R	−1.58857	0.00005	0.01981
TBC1D17	BA10L vs. BA10R	−2.49842	0.00005	0.01981
ATF5	BA10L vs. BA10R	−2.04998	0.00005	0.01981
REXO1	BA10L vs. BA10R	1.50042	0.00005	0.01981
MIR24-2+LOC284454	BA10L vs. BA10R	−4.07299	0.00005	0.01981

Table 1

Gene	Brodmann areas	log2(fold change)	*p–value*	*q–value*
Carcinoembryonic antigen related cell adhesion molecule 1 (Ceacam1), transcript variant 3, mRNA	BA10L vs. BA10R	inf	0.00005	0.01981
Ensemble predicted protein coding	BA10L vs. BA10R	–inf	0.00005	0.01981
HOXD1	BA10L vs. BA10R	–6.65318	0.00005	0.01981
CCNYL1	BA10L vs. BA10R	3.20327	0.00005	0.01981
SLC11A1	BA10L vs. BA10R	2.46311	0.00005	0.01981
NOP56	BA10L vs. BA10R	2.57613	0.00005	0.01981
LAMA5	BA10L vs. BA10R	3.46762	0.00005	0.01981
HMOX1	BA10L vs. BA10R	1.83245	0.00005	0.01981
PISD	BA10L vs. BA10R	2.18628	0.00005	0.01981
PLCL2	BA10L vs. BA10R	3.07933	0.00005	0.01981
TOP2B	BA10L vs. BA10R	–2.91264	0.00005	0.01981
ZBTB20	BA10L vs. BA10R	1.49539	0.0001	0.037021
RUBCN	BA10L vs. BA10R	2.35736	0.00005	0.01981
ARHGAP24	BA10L vs. BA10R	3.26243	0.00005	0.01981
MIR7978	BA10L vs. BA10R	–2.39482	0.00005	0.01981
Non-protein coding gene or gene fragment, no protein prediction available	BA10L vs. BA10R	1.85045	0.00005	0.01981
HSPA1A	BA10L vs. BA10R	3.41616	0.00005	0.01981
AQP1	BA10L vs. BA10R	2.56426	0.00005	0.01981
HSPB1	BA10L vs. BA10R	1.81864	0.00005	0.01981
PDK4	BA10L vs. BA10R	1.49654	0.00005	0.01981
FAM83H	BA10L vs. BA10R	5.24147	0.00005	0.01981
DGAT1	BA10L vs. BA10R	2.09862	0.00005	0.01981
GADD45G	BA10L vs. BA10R	1.94797	0.00005	0.01981
MIR6724-4	BA10L vs. BA10R	3.2094	0.00005	0.01981
STS	BA10L vs. BA10R	–1.8016	0.00005	0.01981
TSIX	BA10L vs. BA10R	–1.96152	0.00005	0.01981
MIR221	BA10L vs. BA10R	–4.19108	0.00005	0.01981
Non-protein coding gene or gene fragment, no protein prediction available	BA10L vs. BA10R	–6.11342	0.00005	0.01981

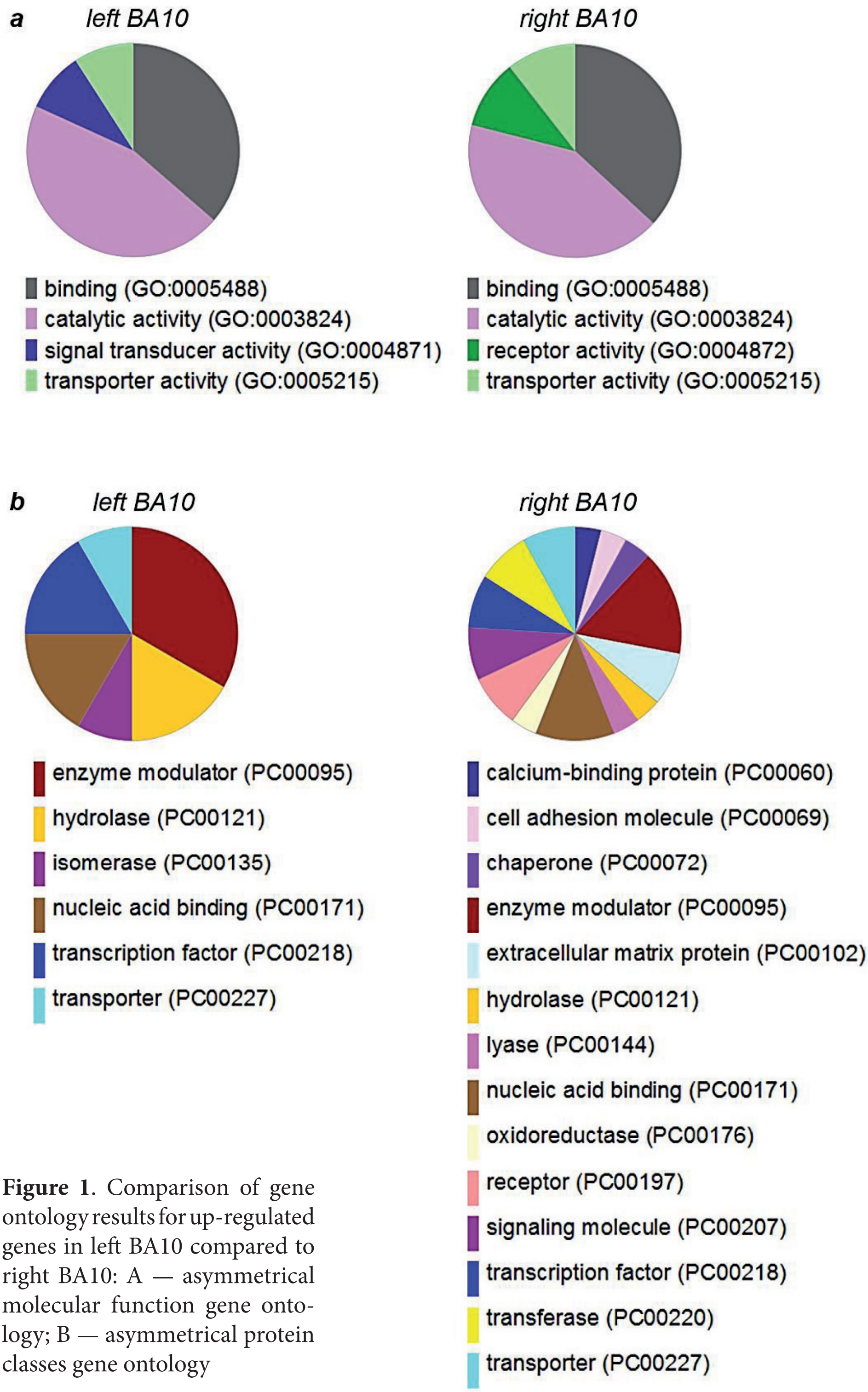


Figure 1. Comparison of gene ontology results for up-regulated genes in left BA10 compared to right BA10: A — asymmetrical molecular function gene ontology; B — asymmetrical protein classes gene ontology

The GO analysis of the gene expression data provides organized terms to describe characteristics of gene products in three categories: biological processes, molecular functions, and cellular components. This analysis also revealed a significant lateralization between left and right PFC. Sharing functions of up-regulated genes — binding, transporter and catalytic activities — were complemented by signal transducer activity in BA10L and by receptor activity in BA10R (Figure 1A). GO analysis of protein classes demonstrated strong increase in diversity of protein classes up-regulated in BA10R compared to BA10L, extending the latter with calcium binding proteins, cell adhesion molecules, receptors, chaperones, extracellular matrix proteins and other classes known to be involved in neuronal plasticity (Baucum, 2017; Dzyubenko, Gottschling, & Faissner, 2016; Gyurko, Soti, Stetak, & Csermely, 2014; Sheng, Leshchyns'ka, & Sytnyk 2013) (Figure 1B).

Further analysis of annotated protein expression in different cell types demonstrated the localization of revealed gene products in the neurons and neuropil. Among 61 studied genes, we found available information on annotated protein expression for 34 genes, with 31 genes showing neuronal expression and with 3 genes having expression in astrocytes and endothelial cells that regulate neuronal excitability and cerebral blood flow according to the activity of synapses (Bazargani, & Attwell, 2016; Nedoluzhko et al., in preparation).

To determine the functional significance of revealed differences between BA10L and BA10R, we also analyzed the data for the possible involvement of genes with differential expression in brain plasticity and activity-dependent mechanisms. In the group of neuronal genes with a strongly lateralized expression, it is worth noting *C-FOS* and *NPAS4* (Figure 2A, 2B, respectively). Both genes show rapid experience-dependent increase in expression levels and are widely used as markers of neuronal plasticity in functional molecular brain map-

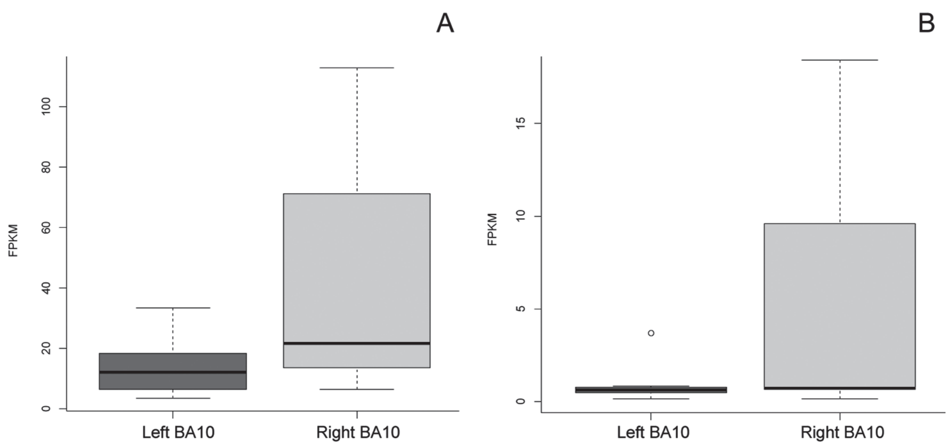


Figure 2. Examples of the relative increase in gene expression in left BA10 compared to right BA10: A — asymmetrical *C-FOS* expression in BA10; B — asymmetrical *NPAS4* expression in BA10

ping in mice (Clayton, 2000; Ivashkina, Toropova, Ivanov, Chekhov, & Anokhin, 2016; X. Sun & Lin, 2016), and also in primate and human brain studies (Caston-Balderrama, Cameron, & Hoffman, 1998; Kaczmarek, Zangenehpour, & Chaudhuri, 1999; Nakagami, Watakabe, & Yamamori, 2013; Okuno & Miyashita, 1996; Rakhade et al., 2005). It is also necessary to note that two other genes, which are particularly strongly up-regulated in the right FPC, namely *SERPINA3* and *mir-331*, are known for their role in molecular mechanisms of hippocampal formation. *SERPINA3* is an endothelial gene having a neurotrophic effect on hippocampal neurons (Kanai, Tanaka, & Hirai, 1991). *Mir-33* is known to determine the expression of the *neuropilin-2* gene (Epis, Giles, Candy, Webster, & Leedman, 2014), a transmembrane receptor gene that regulates the dendritic spine density of pyramidal neurons (Demyanenko et al., 2014). Loss of *neuropilin-2* may induce aberrant processing within hippocampal and corticostriatal networks and thus contribute to neurodevelopmental disorders of memory and motor functions (Shiflett, Gavin, & Tran, 2015).

In the group of 21 genes with expression in the left FPC, we found only two genes with known functionality. One of them is the *miR-34a*, which inhibits expression of its prime target, *sirtuin-1* (*SIRT1*), known to enhance cognitive abilities through proteostatic and neurotrophic mechanisms (Corpas et al., 2016, Lin, Mao, Song, & Huang, 2015). *ATF5*, a repressor of cyclic *AMP* induced transcription (Pati, Meistrich, & Plon, 1999), is also up-regulated in BA10L compared with BA10R, whereas *cAMP* its molecular cascades have been shown to be central in regulating long-term memory and synaptic plasticity (Sheng, Leshchyns'ka, & Sytnyk, 2013).

By analyzing the published databases, we found information on induction-dependent gene expression for 47 of 61 genes in our dataset. Most of them (36 genes) showed inducible character of expression, whereas 11 genes expressed constitutively. There is a second approach to find inducible genes, which is to analyze the published database, consisting of 3804 human housekeeping genes, i.e. genes expected to be expressed irrespective of external signals (see Eisenberg, & Levanon, 2013). Among 21 genes up-regulated in BA10L, the database included four genes considered to be housekeeping (*HAUS4*, *ARF5*, *TOP2B*, *IPO9*). Among 40 genes with higher expression in BA10R, 5 genes were in the list of housekeeping genes (*NSRP1*, *REXO1*, *NFATC2IP*, *NOP56*, *GADD45G*). Thus, according to the second approach, up to 52 genes from the dataset of 61 differentially-expressed genes could be inducible.

Enrichment analysis of discovered genes with differential expression resulted in the detection of 30 genes implicated in mental disorders, and 15 genes related to other conditions, including non-mental brain disorders (gliomas, motor ataxia and others), with 16 genes omitted due to general absence of information on their role in the following categories: neurodevelopmental disorders (*TOP2B, TAOK1, PISD*), major depression (*mir-34A, SPA1A, ARHGAP24*), bipolar disorder (*ATF5, SERPINA3*), autism (*STS, IFITM3*), psychogenic stress (*GAS5, SLC14A1*), Alzheimer disease (*PDK4, SLC11A1*), Parkinson disease (*RGS1*), attention deficit hyperactivity disorder (*STS*) and frontotemporal dementia (*NOP56*). Strikingly, the vast majority of these differentially-expressed genes were implicated in pathogenesis of one single disorder: 24 out of 30 genes were disrupted in schizophrenia. These genes include *mir-34A*, *mir-135B*, *ATF5*, *TOP2B*, *STS*, *SERPINA3*, *HSPA1A*, *RGS1*, *SLC11A1*, *IFITM3*, *MT1X, NSRP1, HSPB1, ZBTB20, C-FOS, BAG3, PDK4, PISD, ARHGAP24, PLCL2, ADM, GADD45G, CD44* and *AQP1*. The list of schizophrenia-related genes demonstrates a rightward asymmetry: 16 up-expressions in BA10R vs. only 8 in BA10L.

Discussion

In the present study, we applied the RNA-Seq technique to analyze lateralization of the molecular mechanisms in FPC, which have previously shown a diversity in anatomical and physiological features between left and right hemispheres, implicated in higher forms of cognitive processes, self-consciousness and voluntary control of behavior. For the first time, to our knowledge, a pronounced difference in molecular mechanisms of BA10L and BA10R was revealed. Until recently, it was often assumed that gene expression in the cerebral cortex is bilaterally symmetrical (Hawrylycz et al., 2012; Pletikos et al., 2014). A new meta-analysis, which was conducted on microarray data, supposed that there are several examples of gene expression lateralization in the superior temporal cortex and auditory cortex of human adults (Karlebach, & Francks, 2015). However, these studies were based on only a small number of postmortem samples limited to brain areas with an obvious functional difference, namely language and speech processing.

Most genes with differential expression in BA10L and BA10R were found to be involved in brain plasticity and activity-dependent molecular mechanisms. Since we analyzed postnatal brain samples with a lack of information on functional cognitive load and with a broad range in postmortal interval, we assume that differential expression of inducible genes primarily reflects re-entry integrative mechanisms and remote cognitive processing (Edelman, & Gally, 2013) akin to activities within the Default Mode Network (DMN). For example, it is known that a baseline expression of immediate early genes is critical for off-line processing of cognitive information (Katche et al., 2010; Makino, Funayama, & Ikegaya, 2016). Moreover, the discovered dominance of inducible genes (about 80%) in our dataset of differential expression allows us to hypothesize that there is a hemispheric lateralization of the DMN in FPC. Recently, a similar lateralization of the effective connectivity has been discovered for a group of key DMN structures by extending dynamic causal modelling to hippocampal formation (Ushakov et al., 2016; Velichkovsky, Krotkova, Sharaev, & Ushakov, 2017).

About a half of the genes with lateralized expression in BA10 were related to mental disorders, wherein the vast majority (80%) were found to be disrupted in schizophrenia. Most of these schizophrenia-related genes (~67%) were up-expressed in BA10R. It can be assumed that FPC has a key involvement in this disease. No such wide association with other mental disorders was found. Schizophrenia is characterized by reduced hemispheric asymmetry of functional brain networks, as shown by recent connectome studies (Y. Sun et al., 2017). On the anatomical level, a similar conclusion was recently made with respect to autism spectrum disorder (Carper, Treiber, DeJesus, & Muller, 2016). Thus, lateralized expression of revealed genes seems to be critical for normal brain functioning.

To date, there is only one comparison of gene expression performed in brain tissue separately from the left and right hemispher areas (Mladinov et al., 2016). These authors studied dorsolateral prefrontal cortex (DLPFC, BA46) and the medial part of the orbitofrontal cortex (MOFC, BA11/12). The only difference they found in normal subjects in the right hemisphere compared with the left hemisphere was in BA11/12, which showed an increased expression of the *KAT7* gene and a decreased expression of gene *NONO*. *KAT7* is a lysine acetyltransferase 7, which may be involved in neuronal plasticity (Feng et al., 2016; Singh & Thakur, 2017). *NONO* is a protein phosphatase 1, regulatory subunit 114, required for tuning inhibitory synapses (Mircsof et al., 2015). In contrast, we revealed 61 genes differential expressed between right and left BA10. The difference in results between these studies could be explained by two reasons. Firstly, there is a long postmortem delay (PMD) time in the study by Mladinov et al. (2016). Samples in their study had ~19 h mean PMD, whereas samples we used had ~12 h mean PMD. As most genes with differential expression are localized to neurons, and neuronal RNA degrades faster than other brain RNAs, we suggest that the shorter PMD in our study could be the critical issue for sensitivity of RNA detection. A second reason could be simply that transcriptomic differences are more pronounced in BA10 than in BA11/12 or BA46 due to differential roles of the left and right PFC (Craik et al., 1999; Grady, Luk, Craik, & Bialystok, 2015; Stuss, & Benson, 1986). This functional lateralization may be reflected in the expression of molecular mechanisms reported in the present study as well as in the Yakovlevian Torque phenomenon known for a long time (LeMay, 1999; Toga, & Thompson, 2003).

On a more conceptual level, the ROBBIA model (ROtman-Baycrest Battery to Investigate Attention) seems to be relevant (Stuss, & Alexander, 2007; Ambrosini, & Vallesi, 2016). The model proposes a prefrontal specialization of two distinct executive functions. One is the "left-lateralized" task-setting function, defined as the transient cognitive control needed to form task-relevant rules and suppress task-irrelevant operations. The other specialization is the "right-lateralized" monitoring function, which provides the cognitive control needed to actively maintain abstract representations by monitoring their status in relation to each other and to intended plans of behaviour. If the right BA10 is in a state of sustained activation in contrast to phasic interventions of the task-setting processes in the left FPC, then we can expect enhanced rightwards asymmetry in activity-dependent molecular mechanisms since they are a target and a probable instrument for enduring maintenance of representations. On the other hand, a relative (48% vs. 12%) and absolute (10 vs. 5 out of 21 and 40) predominance of the non-coding RNA (i.e. expression suppressors) in BA10L would be appropriate to support sporadic acts of selecting and changing domain-specific goals. Indeed, to succeed in multiple choice and selective attention, one needs to suppress task-irrelevant operations. Of course, our hypothesis about contrasting lateralization of protein-coding and regulatory molecular mechanisms has to be tested in direct studies of the microRNA expression within BA10L and BA10R (Nedoluzhko et al., in preparation).

As mentioned above, there is a similarity between this dissociation in FPC and lateralization of processes within DMN. Hippocampal formation and possibly amygdalae are of primary interest here in view of growing evidence for asymmetries in their functional and effective connections (Kerestes, Chase, Phillips, Ladouceur, & Eickhoff, 2017; Ushakov et al., 2016). Although these asymmetries seem to be a feature of the specifically human large-scale brain architecture, their molecular precursors may well be considered in animal study data such as a newfound left-right dissociation of hippocampal memory processes in the mouse (El-Gaby, Shipton, & Paulsen, 2015).

Conclusion

Here, we present results of the first transcriptomic analysis of differential gene expression in the left and right human FPC. A coherent picture of differences is reported, revealing 40 genes that are up-expressed in BA10R and 21 genes up-expressed in BA10L. Differential expression is not confined to the number of genes but to their specialization as well. This has been shown by an additional GO analysis whereby protein-coding genes were predominantly expressed in BA10R and demonstrated a strong increase in diversity of protein classes compared to BA10L. In contrast, a relatively large proportion of up-expressed non-coding RNA has been discovered in BA10L. The results of this study are also potentially of clinical relevance since about half of the discovered genes with lateralized expression in FPC are implicated in mental disorders, first of all in schizophrenia.

Our analysis also opens up several lines of future work. First of all, the sensitivity of such studies could be improved by changing its current design to sequence different brain regions of the same donors. It would be of interest to compare current data from the right FPC with those of the left occipital lobe and in this way possibly contribute to a solution of the Yakovlevian Torque riddle. Even more important is to produce molecular portrayals of the large-scale networks such as DMN with a focus on molecular mechanisms of the amygdalae and both hippocampi which began to be investigated from the point of view of their functional and effective connections ((Kerestes et al., 2017; Ushakov et al. 2016). We hope that future research of gene regulation (e.g. epigenetic studies) at a higher resolution and with a better understanding of large-scale networks will benefit from these early insights into the molecular base of the most anterior parts of the human brain.

Acknowledgements

This work was in part supported by the Russian Science Foundation (grants 14-28-00234 and 17-78-30029 aimed at investigation of self-referential cognition and semantic representations in the human brain, respectively).

References

Ambrosini, E., & Vallesi, A. (2016). Asymmetry in prefrontal resting-state EEG spectral power underlies individual differences in phasic and sustained cognitive control. *Neuroimage, 124*, 843–857. doi: 10.1016/j.neuroimage.2015.09.035

Baucum, A.J., 2nd. (2017). Proteomic analysis of postsynaptic protein complexes underlying neuronal plasticity. *ACS Chemical Neuroscience, 8,* 689–701. doi: 10.1021/acschemneuro.7b00008

Bazargani, N., & Attwell, D. (2016). Astrocyte calcium signaling: The third wave. *Nature Neuroscience, 19,* 182–189. doi: 10.1038/nn.4201

Braver, T.S., & Bongiolatti, S.R. (2002). The role of frontopolar cortex in subgoal processing during working memory. *Neuroimage, 15,* 523–536. doi: 10.1006/nimg.2001.1019

Burgess, P.W., Dumontheil, I., & Gilbert, S.J. (2007). The gateway hypothesis of rostral prefrontal cortex (area 10) function. *Trends in Cognitive Sciences, 11,* 290–298. doi: 10.1016/j.tics.2007.05.004

Carper, R.A., Treiber, J.M., DeJesus, S.Y., & Muller, R.A. (2016). Reduced hemispheric asymmetry of white matter microstructure in autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 55,* 1073–1080. doi: 10.1016/j.jaac.2016.09.491

Caston-Balderrama, A.L., Cameron, J.L., & Hoffman, G.E. (1998). Immunocytochemical localization of Fos in perfused nonhuman primate brain tissue: Fixation and antisera selection. *Journal of Histochemistry and Cytochemistry, 46,* 547–556. doi: 10.1177/002215549804600416

Christoff, K., & Gabrieli, J.D.E. (2000). The frontopolar cortex and human cognition: Evidence for a rostrocaudal hierarchical organization within the human prefrontal cortex. *Psychobiology, 28,* 168–186.

Clayton, D.F. (2000). The genomic action potential. *Neurobiology of Learning Memory, 74,* 185–216. doi: 10.1006/nlme.2000.3967

Colantuoni, C., Lipska, B.K., Ye, T., Hyde, T.M., Tao, R., Leek, J.T., Colantuoni, E.A., … Kleinman J.E. (2011). Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature, 478,* 519–523. doi: 10.1038/nature10524

Corpas, R., Revilla, S., Ursulet, S., Castro-Freire, M., Kaliman, P., Petegnief, V., Gimenez-Llort, M., … Sanfeliu, C. (2016). SIRT1 overexpression in mouse hippocampus induces cognitive enhancement through proteostatic and neurotrophic mechanisms. *Molecular Neurobiology.* doi: 10.1007/s12035-016-0087-9

Craik, F.I.M., Moroz, T.M., Moscovitch, M., Stuss, D.T., Winocur, G., Tulving, E., & Kapur, S. (1999). In search of the self: A positron emission tomography study. *Psychological Science, 10,* 26–34. doi: 10.1111/1467-9280.00102

Demyanenko, G.P, Mohan, V., Zhang, X., Brennaman, L.H., Dharbal, K.E., Tran, T.S., … Maness, P.F. (2014). Neural cell adhesion molecule NrCAM regulates Semaphorin 3F-induced dendritic spine remodeling. *Journal of Neuroscience, 34,* 11274–11287. doi: 10.1523/JNEUROSCI.1774-14.2014

Dzyubenko, E., Gottschling, C., & Faissner, A. (2016). Neuron-glia interactions in neural plasticity: Contributions of neural extracellular matrix and perineuronal nets. *Neural Plasticity, 2016,* 5214961. doi: 10.1155/2016/5214961

Edelman, G.M., & Gally, J.A. (2013). Reentry: A key mechanism for integration of brain function. *Frontiers in Integrative Neuroscience, 7,* 63. doi: 10.3389/fnint.2013.00063

Eisenberg, E., & Levanon, E.Y. (2013). Human housekeeping genes, revisited. *Trends in Genetics, 29,* 569–574. doi: 10.1016/j.tig.2013.05.010

El-Gaby M., Shipton, O.A., & Paulsen, O. (2015). Synaptic plasticity and memory: New insights from hippocampal left-right asymmetries. *Neuroscientist*, *21*(5), 490–502. doi: 10.1177/1073858414550658.

Epis, M.R., Giles, K.M., Candy, P.A., Webster, R.J., & Leedman, P.J. (2014). miR-331-3p regulates expression of neuropilin-2 in glioblastoma. *Journal of Neuro-oncology, 116,* 67-75. doi: 10.1007/s11060-013-1271-7

Feng, Y., Vlassis, A., Roques, C., Lalonde, M.E., Gonzalez-Aguilera, C., Lambert, J.P., … Groth, A. (2016). BRPF3-HBO1 regulates replication origin activation and histone H3K14 acetylation. *The EMBO Journal, 35,* 176–192. doi: 10.15252/embj.201591293

Forgács, B., Lukács, Á., & Pléh, C. (2014). Lateralized processing of novel metaphors: Disentangling figurativeness and novelty. *Neuropsychologia, 56,* 101–109. doi: 10.1016/j.neuropsychologia.2014.01.003

Grady, C.L., Luk, G., Craik, F.I., & Bialystok, E. (2015). Brain network activity in monolingual and bilingual older adults. *Neuropsychologia, 66,* 170–181. doi: 10.1016/j.neuropsychologia.2014.10.042

Gyurko, D.M., Soti, C., Stetak, A., & Csermely, P. (2014). System level mechanisms of adaptation, learning, memory formation and evolvability: The role of chaperone and other networks. *Currunt Protein and Peptide Science, 15,* 171–188. doi: 10.2174/1389203715666140331110522

Halpern, M.E., Gunturkun, O., Hopkins, W.D., & Rogers, L.J. (2005). Lateralization of the vertebrate brain: Taking the side of model systems. *Journal of Neuroscience, 25,* 10351–10357. doi: 10.1523/JNEUROSCI.3439-05.2005

Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, … Jones AR. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature, 489,* 391–399. doi: 10.1038/nature11405

Herbert, M.R., Ziegler, D.A., Deutsch, C.K., O'Brien, L.M., Kennedy, D.N., Filipek, P.A., … Caviness, V.S.Jr., (2005). Brain asymmetries in autism and developmental language disorder: A nested whole-brain analysis. *Brain, 128,* 213–226. doi: 10.1093/brain/awh330

Hrvoj-Mihic, B., Bienvenu, T., Stefanacci, L., Muotri, A.R., & Semendeferi, K. (2013). Evolution, development, and plasticity of the human brain: from molecules to bones. *Frontiers in Human Neuroscience, 7,* 707. doi: 10.3389/fnhum.2013.00707

Huang da, W., Sherman, B.T., & Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols, 4,* 44–57.

Ivashkina, O.I., Toropova, K.A., Ivanov, A.A., Chekhov, S.A., & Anokhin, K.V. (2016). Waves of c-Fos and Arc Ppoteins expression in neuronal populations of the hippocampus in response to a single episode of new experience. *Bulletin of Experimental Biology and Medicine, 160,* 729–732. doi: 10.1007/s10517-016-3296-3

Jia, P., Han, G., Zhao, J., Lu, P., & Zhao, Z. (2017). SZGR 2.0: A one-stop shop of schizophrenia candidate genes. *Nucleic Acids Research, 45,* D915-D924. doi: 10.1093/nar/gkw902

Kaczmarek, L., Zangenehpour, S., & Chaudhuri, A. (1999). Sensory regulation of immediate-early genes c-fos and zif268 in monkey visual cortex at birth and throughout the critical period. *Cerebral Cortex, 9,* 179–187. doi: 10.1093/cercor/9.2.179

Kanai, H., Tanaka, M., & Hirai, S. (1991). Alpha-1-Antichymotrypsin has a trophic effect on hippocampal-neurons invitro. *Neuroscience Letters, 125,* 163–165. doi: 10.1016/0304-3940-(91)90017-N

Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., … Sestan, N. (2011). Spatio-temporal transcriptome of the human brain. *Nature, 478,* 483–489. doi: 10.1038/nature10523

Karlebach, G., & Francks, C. (2015). Lateralization of gene expression in human language cortex. *Cortex, 67,* 30–36. doi: 10.1016/j.cortex.2015.03.003

Katche, C., Bekinschtein, P., Slipczuk, L., Goldin, A., Izquierdo, I.A., Cammarota, M., & Medina, J.H. (2010). Delayed wave of c-Fos expression in the dorsal hippocampus involved specifically in persistence of long-term memory storage. *Proceedings of the National Academy of Sciences of the United States of America, 107,* 349–354. doi: 10.1073/pnas.0912931107

Kerestes, R., Chase, H.W., Phillips, M.L., Ladouceur, C.D., & Eickhoff, S.B. (2017). Multimodal evaluation of the amygdala's functional connectivity. *Neuroimage, 148,* 219–229. doi: 10.1016/j.neuroimage.2016.12.023

LeMay, M. (1999). Functional and anatomical asymmetries of the human brain. *European Journal of Neurology, 6,* 79–85. doi: 10.1046/j.1468-1331.1999.610079.x

Lin, Q., Mao, Y.R., Song, Y.L., & Huang, D.F. (2015). MicroRNA-34a induces apoptosis in PC12 cells by reducing B-cell lymphoma 2 and sirtuin-1 expression. *Molecular Medicine Reports, 12,* 5709–5714. doi: 10.3892/mmr.2015.4185

Makino, K., Funayama, K., & Ikegaya, Y. (2016). Spatial clusters of constitutively active neurons in mouse visual cortex. *Anatomical Science International, 91,* 188–195. doi: 10.1007/s12565-015-0284-z

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., & Thomas, P.D. (2017). PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Research, 45,* D183-D189. doi: 10.1093/nar/gkw1138

Miller, E.K., & Cohen, J.D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167–202. doi: 10.1146/annurev.neuro.24.1.167

Mircsof, D., Langouet, M., Rio, M., Moutton, S., Siquier-Pernet, K., Bole-Feysot, C., … Colleaux L. (2015). Mutations in NONO lead to syndromic intellectual disability and inhibitory synaptic defects. *Nature Neuroscience, 18,* 1731–1736. doi: 10.1038/nn.4169

Mladinov, M., Sedmak, G., Fuller, H.R., Babic Leko, M., Mayer, D., Kirincich, J., … Simic, G. (2016). Gene expression profiling of the dorsolateral and medial orbitofrontal cortex in schizophrenia. *Translational Neuroscience, 7,* 139–150. doi: 10.1515/tnsci-2016-0021

Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods, 5,* 621–628. doi: 10.1038/nmeth.1226

Nakagami, Y., Watakabe, A., & Yamamori, T. (2013). Monocular inhibition reveals temporal and spatial changes in gene expression in the primary visual cortex of marmoset. *Frontiers in Neural Circuits, 7,* 43. doi: 10.3389/fncir.2013.00043

Nedoluzhko, A.V., Efimova, O.I., Kildyushov, E.M., Dolina, I.A., Sokolov, A.S., Ushakov, V.L., Khaitovich, P.E. … Velichkovsky, B.M. (in preparation). Sequencing frontopolar cortex: New insights into the molecular base of human cerebral asymmetry.

Okuno, H., & Miyashita, Y. (1996). Expression of the transcription factor Zif268 in the temporal cortex of monkeys during visual paired associate learning. *European Journal of Neuroscience, 8,* 2118–2128. doi: 10.1111/j.1460-9568.1996.tb00733.x

Pati, D., Meistrich, M.L., & Plon, S.E. (1999). Human Cdc34 and Rad6B ubiquitin-conjugating enzymes target repressors of cyclic AMP-induced transcription for proteolysis. *Molecular and Cellular Biology, 19,* 5001–5013. doi: 10.1128/MCB.19.7.5001

Penfield, W., & Evans, J. (1935). The frontal lobe in man: A clinical study of maximum removals. *Brain, 58,* 115–133. doi: 10.1093/brain/58.1.115

Pletikos, M., Sousa, A.M., Sedmak, G., Meyer, K.A., Zhu, Y., Cheng, F. … Sestan, N. (2014). Temporal specification and bilaterality of human neocortical topographic gene expression. *Neuron, 81,* 321–332. doi: 10.1016/j.neuron.2013.11.018

Ponten, F., Gry, M., Fagerberg, L., Lundberg, E., Asplund, A., Berglund, L., … Uhlen, M. (2009). A global view of protein expression in human cells, tissues, and organs. *Molecular Systems Biology, 5,* 337. doi: 10.1038/msb.2009.93

Pujol, J., Lopez-Sala, A., Deus, J., Cardoner, N., Sebastian-Galles, N., Conesa, G., & Capdevila, A. (2002). The lateral asymmetry of the human brain studied by volumetric magnetic resonance imaging. *Neuroimage, 17*, 670–679. doi: 10.1006/nimg.2002.1203

Rakhade, S.N., Yao, B., Ahmed, S., Asano, E., Beaumont, T.L., Shah, A.K., … Loeb, J.A. (2005). A common pattern of persistent gene activation in human neocortical epileptic foci. *Annals Neurology, 58,* 736–747. doi: 10.1002/ana.20633

Ramnani, N., & Owen, A.M. (2004). Anterior prefrontal cortex: Insights into function from anatomy and neuroimaging. *Nature Reviews Neuroscience, 5,* 184–194. doi: 10.1038/nrn1343

Renteria, M.E. (2012). Cerebral asymmetry: A quantitative, multifactorial, and plastic brain phenotype. *Twin Research and Human Genetics, 15,* 401–413. doi: 10.1017/thg.2012.13

Semendeferi, K., Armstrong, E., Schleicher, A., Zilles, K., & Van Hoesen, G.W. (2001). Prefrontal cortex in humans and apes: A comparative study of area 10. *American Journal of Physical Anthropology, 114,* 224–241. doi: 10.1002/1096-8644(200103)114:3<224::AID-AJPA1022>3.0.CO;2-I

Sheng, L., Leshchynska, I., & Sytnyk, V. (2013). Cell adhesion and intracellular calcium signaling in neurons. *Cell Communication and Signaling, 11,* 94. doi: 10.1186/1478-811X-11-94.

Shiflett, M.W., Gavin, M., & Tran, T.S. (2015). Altered hippocampal-dependent memory and motor function in neuropilin 2-deficient mice. *Translational Psychiatry, 5,* e521. doi: 10.1038/tp.2015.17.

Singh, P., & Thakur, M.K. (2017). Histone deacetylase 2 inhibition attenuates downregulation of hippocampal plasticity gene expression during aging. *Molecular Neurobiology*. doi: 10.1007/s12035-017-0490-x

Soch, J., Deserno, L., Assmann, A., Barman, A., Walter, H., Richardson-Klavehn, A., & Schott, B.H. (2016). Inhibition of information flow to the default mode network during self-reference versus reference to others. *Cerebral Cortex.* doi: 10.1093/cercor/bhw206

Stuss, D.T., & Alexander, M.P. (2007). Is there a dysexecutive syndrome? *Philosophical Transactions of the Royal Society B, 362,* 901–915. doi: 10.1098/rstb.2007.2096

Stuss, D.T., & Benson, D.F. (1986). *The frontal lobes*. New York: Raven Press.

Sun, T., Collura, R.V., Ruvolo, M., Walsh, & C.A. (2006). Genomic and evolutionary analyses of asymmetrically expressed genes in human fetal left and right cerebral cortex. *Cerebral Cortex, 16 Suppl 1,* i18–25. doi: 10.1093/cercor/bhk026

Sun, X., & Lin, Y. (2016). Npas4: Linking neuronal activity to memory. *Trends in Neurosciences, 39,* 264–275. doi: 10.1016/j.tins.2016.02.003

Sun, Y., Chen, Y., Collinson, S.L., Bezerianos, A., & Sim, K. (2017). Reduced hemispheric asymmetry of brain anatomical networks is linked to schizophrenia: A connectome study. *Cerebral Cortex, 27,* 602–615.

Toga, A.W., & Thompson, P.M. (2003). Mapping brain asymmetry. *Nature Reviews Neuroscience, 4,* 37–48. doi: 10.1038/nrn1009

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, … Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols, 7*, 562–578. doi: 10.1038/nprot.2012.016

Ushakov, V.L., Sharaev, M.G., Kartashov, S.I., Zayyalova, V.V., Verkhlyutov, V.M., & Velichkoysky, B.M. (2016). Dynamic Causal Modeling of hippocampal links within the Human Default Mode network: Lateralization and computational stability of effective connections. *Frontiers in Human Neuroscience, 10,* 528. doi: 10.3389/fnhum.2016.00528

Velichkovsky, B.M., Krotkova, O.A., Sharaev, M.G., & Ushakov, V.L. (2017). In search of the "I": Neuropsychology of lateralized thinking meets Dynamic Causal Modeling. *Psychology in Russia: State of the Art, 10*(3), 7–27. doi: 10.11621/pir.2017.0301

Original manuscript received August 25, 2017
Revised manuscript accepted September 11, 2017
First published online September 30, 2017