A New Approach to the Study of Consciousness from the Theory of Integrated Information Point of View (Review)

DOI: 10.17691/stm2016.8.4.30
Received April 6, 2016

L.A. Legostaeva, PhD Student, Neurologist, Neurocritical Care;
E.A. Zmeykina, Research-Engineer, Neuro-Rehabilitation Department;
A.G. Poydasheva, Junior Researcher, Neurologist, Neurophysiologist, Neuro-Rehabilitation Department;
D.V. Sergeev, MD, PhD, Neurologist, Department of Neurocritical Care;
D.O. Sinitsyn, PhD, Engineer-Researcher, Neuro-Rehabilitation Department;
E.I. Kremneva, MD, PhD, Senior Researcher, Radiologist, Department of Neuroimaging;
A.V. Chervyakov, MD, PhD, Senior Researcher, Neurologist, Neurophysiologist, Neuro-Rehabilitation Department;
Yu.V. Ryabinkina, MD, DSc, Leading Researcher, Department of Neurocritical Care;
N.A. Suponeva, MD, PhD, Neurologist, Neurophysiologist, Head of the Neuro-Rehabilitation Department;
M.A. Piradov, MD, DSc, Professor, Academician of the Russian Academy of Sciences, Director
Research Centre of Neurology, 80 Volokolamskoye Shosse, Moscow, 125367, Russian Federation

The continuously increasing number of patients with chronic disorders of consciousness results in a growing need for an objective method of evaluating the level of consciousness in this category of patients in the clinical practice of neurologists and critical care physicians. One of the up-to-date approaches is based on the integrated information theory (IITC; Tononi, 2004, 2012, 2014). According to this theory the mechanisms determining the state of consciousness are integrally connected with certain conceptually coordinated structures that are closely interrelated with each other, and exist independently of external influences. On the basis of this theory Casali et al. (2013) developed approaches to the quantitative evaluation of integrated information in the neuronal system that depend on cause-and-effect relationships between its different elements. As a result, the term ‘perturbational complexity index’ has been introduced in respect of consciousness. It allows measurement of the integrity and spatio-temporal structure of the pattern of cortical excitation caused by non-invasive stimulation of the brain cortex, and, therefore, a description of the degree of complexity of the electrical signals in the brain. The values of this index differ in various states of consciousness (e.g., wakefulness, deep sleep, anesthesia, vegetative state etc.) that supports its use in patients with chronic disorders of consciousness as a new diagnostic tool.

Key words: functioning of consciousness; minimally conscious state; vegetative state; integrated information theory; neuronal system; transcranial magnetic stimulation.

Introduction. From the point of view of a clinical neurologist or critical care physician the issue of studying consciousness is mainly relevant to fundamental studies in the field of neuroscience. However, professionals in these two specialties are often faced with phenomena of impaired consciousness in their practice [1]. The determination of the presence or absence of conscious activity in patients in a vegetative state (also known as the unresponsive wakefulness syndrome) or in a minimally conscious state can be decisive in a prognosis of the recovery of brain functions. Despite the clear diagnostic criteria of different variants of chronic disorders of consciousness (DOC) [2], their accurate diagnosis is a complex clinical problem. Currently, the clinical evaluation of a patient’s status is based on examining their behavioral response to external stimuli and the subjective interpretation of these by the physician, but this approach can result in a high frequency of diagnostic mistakes — reaching 37–43% [3–7]. To avoid this, various attempts to evaluate consciousness with neurovisualization and neurophysiological approaches are being actively developed [8–15]. A clear understanding of current trends in the field of consciousness studies requires a fundamental approach based on the results of investigating the neuronal correlates of consciousness, i.e. dynamic processes corresponding to episodes of consciousness. Understanding consciousness from the point of view of the modern integrated information theory allows practicing physician to better realize the value of the up-to-date methods in the study of disorders of consciousness [16].

Neurology of consciousness. Classic neurology defines consciousness as a state in which a person is aware of what is happening with him himself and in the surroundings. Consciousness is determined by two
The 20th century saw a shift toward the development of approaches for patients with severe impairments of brain functions, expressed in a lack of any signs of cognitive activity, while the vegetative functions are almost completely preserved [24–28]. In the M斯基 there is reproducible behavioral evidence of self-awareness and of awareness of the surrounding reality [26, 29, 30].

The question of the presence or absence of signs of consciousness in patients cannot, and should not, be based solely on subjective evaluation. It is therefore crucial to determine diagnostic and neuro-rehabilitating approaches for patients with severe impairments of brain of hypoxic or traumatic origin and who are in a chronic DOC state [31, 32]. A study of the neuronal activity occurring in a state of impaired consciousness can be useful for understanding the process of consciousness both from a fundamental point of view and in the research of possible background for neuro-rehabilitation.

**Fundamental approach.** Throughout the history of the mankind philosophical attempts have been made at understanding the phenomenon of consciousness and its relationship to the physical body and the world outside. For example, starting with Plato, ideas on the dualism of body and mind have been repeatedly expressed. In the 20th century there was a shift toward the development of specific, empirically testable theories of mental functioning in humans and animals, including of consciousness.

The opportunity to generate such theories was a result of the rapid accumulation of experimental data on the functioning of the nervous system on the one hand, and on the other, by the development of artificial informational systems and an understanding of the general principles of how to work with information.

Today one of the most interesting areas of neuroscience is the study of the process of consciousness [33, 34]. With the development of computer technology in the 1940s it became possible to model neural networks and this led to a completely new understanding of how to represent neuronal functions through the identification of the significant operating elements and the algorithms of their action [35]. The interest and the originality of this approach, suggested by mathematicians, consisted in the fact that the nervous system could be described as a group of interconnected elements — neurons, and the neural network as a means to summarize any information obtained, transforming it into mathematical functions. Thus, a processor model of brain functioning was proposed. In this model the global processes responsible for the cognitive work of the brain are considered to be the result of the work of neuron ensembles [36, 37]. The presence of a connection between neuron ensembles and the information obtained from the senses questioned the issue of the mechanisms used by the brain for reading and decoding external information.

This system, which presupposes an absence of static processes but the possibility of spontaneous changes in the informational environment, is at the same time both complex and simple. According to Hebb's theory, neurons are trained through changes in the strength of their synaptic connections resulting from frequent stimulation both from outside and inside [38]. This idea combines the classic perception of their physical and chemical properties with new possibilities of complex multi-level changes in the system resulting from interactions (plasticity). The idea arose of new connections that could be combined systematically according to hierarchically performed functions.

A functional and anatomical hierarchy built on principles of brain activity suggests the global nature of its work. However, multidimensional and integrated functions implying pre- and post-processing cannot be realized by a single neuron working as a structural-functional unit of the nervous system. The framework for complex cognitive processes is not just one, or several neurons, but a complex neural network the operation of which cannot be described by a linear equation [39]. Through natural selection it was possible to unite neurons into organized systems which can recognize and process information of different levels of complexity [40–44]. These organized systems of neurons form neuronal maps. Moreover, such an abstract representation of interacting systems, united in networks, can be transformed into a logical hierarchical structure with clear localization and with the specific algorithms involved in the formation of the processes of functioning consciousness. In this context, conventional ideas about the structural-functional system of the brain's operation must be reconsidered, being supplemented with new possibilities and principles of functioning. The idea appears of two-levels of registration, processing, storage and use of the information, with ongoing categorization of the perceived impulses distributing them into stable information hubs. Then these hubs, in their turn, are united into a stable dynamic nucleus [45, 46] that combines information flows from outside (through analyzers) and from inside. Generally, if a significant part of the dynamic nucleus loses its capacity for integrating [47], all its activity is impossible. The integration of the elements and the processed information matches the anatomy of the thalamo-cortical system and determines its complexity and the dynamics of its functioning: conscious behavior depends on the interaction between the activating brain systems [29]. These systems combine...
the cortico-cortical, cortico-thalamic and thalamo-cortical connections which structurally provide dynamic re-entry and excitation loops [48]. The thalamo-cortical connections can synchronize excitation in the brain cortex. This thalamo-cortical system is physically uniform, but includes a large number of functionally-specialized elements which are activated selectively [49]. The result of the division into subsystems means that it is possible to specialize the elements and integrate their functions, allowing rapid and effective interaction both within the small elements of neuron ensembles and within the larger system as a whole [50]. This system organization emphasizes both its functional and dynamic specialization as well as the integration of its individual elements, but such a system requires a high level of complexity and is highly differentiated [47]. Modeling of thalamo-cortical interactions [51, 52] can be represented by a uniform synchronization of rhythmic oscillations (in the gamma-range) over time. Excitation occurs during coordination of the functioning of the two loops of the thalamo-cortical system (basic and specific levels), and this is modulated by the brain stem [53]. On the other hand, large-scale rhythmic fluctuations within the cerebral cortex, reflecting the activity of such global processes as consciousness, can act competitively when a more important process of excitation becomes dominant [54].

The neuronal correlates of consciousness are considered to be associated with the connections between the neurons as these are sufficient for the required information processing [54–58]. This information processing is a key characteristic of the formation of consciousness.

The main approaches to the construction of a quantitative theory of systems possessing consciousness are united by the notions of information quantity [59] and complexity [60]. In the study of the work by G. Tononi et al. [61] a measure of the complexity of the neuronal activity was proposed, this theory being based on the idea of the co-existence of integration and differentiation in the operation of the brain neurons. So, on the one hand, the activity of the brain neurons is fairly differentiated, with single neuron groups obtaining different input signals and having different dynamics of activity. On the other hand, these groups are considerably interconnected at an anatomical level (in essence any two neurons are connected with each other through a chain of just a few synapses, providing a plurality of feedback connections), and at a functional level (there are widespread patterns of activity correlation). In this regard neither the occasional (only differentiation) nor the homogeneous (only integration) patterns of activation have a high degree of complexity. This property only applies to systems with a large number of different patterns in the working of the neuronal connections [51, 62–65].

The theory of integrated information by G. Tononi. A mathematical model of the consciousness process, applied in the clinical and neurophysiological investigation of patients with impaired consciousness, is reflected in the integrated information theory by G. Tononi [45, 66–68]. The primary axiom of this theory is fairly simple: it claims that consciousness exists. Then the author gives a number of statements which form the theoretical basis for a mathematical approach to understanding the process of consciousness.

Consciousness is a complex, multicomponent, structured set of combinations of the states of the brain system and its connections under conditions of continuously gained experience [23]. Such obtained experience is specific and exclusive; it is absolutely different from any other experience. The obtained experience is a single whole that cannot be divided into independent components since they are integrated. The states of the neuronal system change selectively due to the work of certain complex hierarchical mechanisms that can generate information under the influence of effective factors. The mechanisms within these states generate integrated information that cannot be divided into multiple independent parts since they use minimal information space for this generation. Moreover the generation of consciousness is exclusive in terms of time, meaning that at any single moment a single state of consciousness dominates.

The integrated information theory suggests that the mechanisms involved in the state of consciousness are continuously connected with specific conceptually coordinated structures which are united with each other irrespective of the quality of external influence. Each specific mechanism works in accordance with the "cause-and-effect" principle, i.e. its functioning is built as a result of selected alternatives. Thus, the forms of the pairs of cause-and-effect elements are unique and exist only as a result of information integration within a particular state.

The cause of this is the correlation of consecutively observed events that generate a transfer of mutual information from one element to another within the system. At rest, it is impossible to identify the causative element in the system of mechanisms forming such an integrated network of the observed events. However, if the system is excited from the outside, it is possible to build a model of cause-and-effect connections where one of the elements will definitely be an effect, given that the system is unique and integrated [15, 69].

Tononi's work [70] also proposed the concept of determining the integration of information within the system using a search of method that aims to separate it into two parts minimally connected with each other. The resulting value measures the extent to which the system cannot "be further reduced" to the sum of its parts — the theory of consciousness as integrated information is based on this aspect. Thus the suggested theoretical measure of integrated information is based on the cause-and-effect connections between the system elements, rather than on statistical correlation as standard measures of the quantity of information [61, 71, 72].

However, apart from the conceptual counterarguments [73] facing the theory of integrated information, there is the
The technical problem of finding such integrated information: even for model systems containing not more than a dozen elements, an accurate mathematical calculation is almost impossible in practice [68]. This makes direct experimental verification of the theory extremely difficult. However, studies on a definition of an integrated index of consciousness [74, 75] obtained experimental evidence in favor of the idea of the combination of the integration and differentiation of information within the dynamics of the system as a prerequisite for the presence of consciousness.

**Experimental proof of the theory of integrated information.** Consciousness unlike other cognitive functions (such as, for example, memory or image recognition) is characterized by an integral subjective aspect: a cognitive process is conscious, if it is realized internally by the individual and is a part of their subjective experience [76]. External observations do not guarantee detection of this aspect and this causes difficulties in verifying theoretical statements about the consciousness process. In comparison, consider studies of pseudo-blindness [77, 78] where patients with an impaired visual cortex were able to carry out tasks that required vision so accurately that it could not be just a coincidence, despite their report that they did not have any visual stimulus. Study [79] similarly describes experiments testifying to the possibility of performing unconscious word processing and arithmetic calculations. A subjective component of consciousness also underlies the philosophical discussions of this problem [80].

To be eligible for acceptance any theory must be tested empirically. Such important issues as the neuroanatomy, neuropsychology and neurophysiology of consciousness usually stop at the stage of searching for the neuron correlates of consciousness. To determine the role of the cortico-thalamic system and the cortico-cortical interactions in the mechanisms of consciousness as well as the presence of dominant mechanisms within the state of consciousness a cause-and-effect approach to the theory of integrated information can be useful.

Its realization became possible due to the combination of a mathematical approach and clinical studies. Progress in the field of neurophysiological and neurovisualization methods of diagnostics enabled a study involving a comparison of healthy volunteers having clear consciousness to patients with impaired consciousness.

Casali et al. [75] developed a quantitative measure based on the integrated information theory [66] that can provide an indicator of the parameters of consciousness-potential regardless of external sensory stimulus, possible motor response or the direct participation of the patient in the study. This measure was named the perturbational complexity index (PCI). It is calculated by analyzing electrical activity responses, generated by transcranial magnetic stimulation (TMS), and registered with an electroencephalogram (EEG): this parameter measures both the complexity and the spatial-temporal structure of the pattern determined by the cortical excitation resulting from non-invasive stimulation of the cerebral cortex. The physical concept of the index is as an evaluation of the complexity of the electrical signals within the brain. In other words, by measuring electrophysiological activity after TMS it is possible not only to follow the distribution of the signal on the surface of the cerebral cortex, but to indicate the extent of “consciousness” through this distribution. A higher degree of compressibility of the TMS-spatial-temporal pattern of activity presupposes an orderly and “uncomplicated” system, while a lower degree of compressibility indicates the presence of a complex information-rich system of connections in the brain cortex. Thus, it is suggested that in different states of consciousness (for example, in deep sleep, anesthesia, a vegetative state etc.) the index reflecting the complexity of processing an external signal in the brain will decrease compared to the maximum value obtained during a state of wakefulness.

This index is calculated in several stages: 1) the TMS-generated potentials are registered with a 60-channel EEG; 2) solving inverse problem for EEG source of brain electrical activity isolated from the record are localized onto an anatomical image of the brain obtained with an MRI; 3) a spatial-temporal matrix of the sources is binarized with statistical transformations, i.e. SS(x,t)=1 for statistically significant sources, and SS(x,t)=0 for the remainder.

To calculate the PCI index the Lempel–Ziv algorithm is used for compressing the spatial-temporal matrix of the sources [65]. The Lempel–Ziv complexity (c) approximates the nonredundant information contained in the binary matrix of length L by estimating the minimum number of different patterns required to describe the sequence. The asymptomatic behavior of this measure for a random sequence is LH(L)/log2L, where H(L) is the source of entropy

\[ H(L) = -p \log_2(p) - (1-p) \log_2(1-p) \]

and where p is the proportion of units contained in the binary matrix with length L.

The PCI is defined as the normalized Lempel–Ziv complexity of the TMS-generated spatial-temporal pattern of activity of the brain cortex SS(x,t):

\[ \text{PCI} = \frac{\log_2L}{\text{LH}(L)} \]

The preliminary results presented by Casali et al. [75, 81] show that healthy people, while unconscious — in a state of deep sleep or under anesthesia (caused by pharmaceuticals such as midazolam, propofol and xenon) — had a considerably lower PCI when compared to that in a state of wakefulness. Furthermore, for patients in a state of minimum consciousness the PCI was higher compared to that in patients in a vegetative state, and this correlates with the data obtained from patients with “locked-in” syndrome, who, as is known, are in clear consciousness.
Electrophysiological activity generated by transcranial magnetic stimulation: normal (a), in a minimally conscious state (b), in a vegetative state (c).
Thus, the study of consciousness requires an interdisciplinary approach. In particular, the theory of integrated information can be applied to identify conscious activity in patients with chronic DOC (See the Figure). A mathematical calculation of the integrative index of consciousness (PCI) allows not only a description of the degree of complexity of the functioning of the thalamocortical system, but provides an opportunity to empirically evaluate the state of consciousness.

Conclusion. Determination of the presence or absence of signs of conscious activity is a complicated clinical task, but the possibilities of using a mathematical approach to modeling the mechanisms of consciousness, in particular on the basis of the integrated information theory, can provide an objective standard in the determination of signs of conscious activity in patients in a vegetative state or in a minimally conscious state. The work was funded by a grant from the Russian Science Foundation No.16-15-00274 “Consciousness within a framework of integrated information theory”.

Conflicts of interest. The authors declare no conflicts of interest.

References


23. Giacino J.T., Kalmar K., Whyte J. The JFK coma recovery scale-revised: measurement characteristics and
62. Seth A.K., Barrett A.B., Barnett L. Causal density and integrated information as measures of conscious level. Philos


