Consciousness enigma: The "hard problem" – binding problem entanglement, "extra ingredient" and field principle

Michael Lipkind
Unit of Molecular Virology, Kimron Veterinary Institute, Beit Dagan, Israel; International Institute of Biophysics, Neuss-Hombroich, Germany

The "Hard problem" of consciousness relates to the perplexing emergence of consciousness as a result of the brain activity. The binding problem concerns separate processing of perceived data in functionally and topographically segregated cortical areas and functional integration of such locally disjoined operations into coherently perceived images and events. The existing field-grounded theories of consciousness fall into two groups: (i) those based on the physical (electromagnetic) field, and (ii) those grounded on autonomous fields irreducible to the established physical fundamentals. The critical analysis of the existing theories results in formulation of the extra ingredient and a novel irreducible field-based theory of consciousness.

Keywords: Binding problem, Consciousness emergence, Extra ingredient, Field principle, Hard problem, Protoconsciousness, Protophenomenal fundamental

The problem of consciousness takes up an especial place within the frame of the general scientific paradigm. Moreover, there is a conviction that the problem of consciousness transcends purely scientific limits reaching the acme of philosophical sphere. At times, such an "over-generalized" conclusion has had a paralyzing effect on scientific endeavor, and only recently, the consciousness problem has been transformed into the modern trend "Toward a science of consciousness". However, the new transformation has not solved the situation, and the problem of consciousness is still displays insuperable difficulties. Nevertheless, on such a hopeless background, more specific problems, namely, the binding problem and the "hard problem" of consciousness, have been declared, which, in contrast to the vaguely defined and highly complicated Consciousness per se, are clearly formulated.

The binding problem

Historically, the binding problem was connected first with psychology being dependent on general philosophical views on spatial-temporal contiguity of mental representation of the external world. The modern version of the binding problem is expressed on neurological level being based on the well established evidence on the disjunctive way of processing of visible percepts realized within different brain cortex areas. In particular, the visual data are processed separately within about fifty functionally segregated specialized cortical areas, each one being responsible for a specific feature, like movement, colour, texture, size, curvature, some topological properties like height/width ratio, stereoscopic depth, orientation of lines and edges, and so on. In spite of such spatially segregated processing of particular features of the object, it is perceived as unitary, i.e. the whole dynamically changing world appears as integral and coherent. Therefore, the synthesis of all the disjointed, dispersed, and separately processed elements (components, features) of the complex signals from the continually changing picture of the external world must be realized by binding together neurological states occurring in different brain areas. At the same time, multi-modal association areas in the cortex in which single perceptual features could be unified into a final perceptual image have not been found, so there is no explanation how the disjointed features of any perceived object are linked together. Such a "binding problem" looks rather mysterious since there is no locus in the cortex, which could be called either...
“master map”\textsuperscript{13}, or multi-modal association areas\textsuperscript{18}, or central cortical “information exchange”\textsuperscript{19}. Nevertheless, still there is a certain (rather emotional) hope for future finding of “grandmother” neurons and convergence zones\textsuperscript{19}.

The binding problem is particularly demonstrable in the case of the visual system. The histological fact is that about 300 retinal rods (the 1st neuron of the visual network way) are structurally (histologically) connected via bipolar cells (the 2nd neuron) with one ganglion cell (the 3d neuron). Consequently, these 300 adjacent rods form a microarea within retina which during the vision process may contain heterogeneous picture as projected from a micro-part of the visible object. Thus, the axon of the ganglion cell must conduct forward an impulse carrying such complex (integrated) visual information. This already is incompatible with the classic neuronal theory according to which the neuron firing means only the conduction of a signal that is realized according to the ‘all-or-none’ principle.

Thus, apart from the binding problem seemingly realized on the cortical level, two opposite processes associated with the same current visual signal go on simultaneously. One of them consists of the anatomically determined confluence of distinct signals related to different spatial parts of the perceived object into a single but complex signal, its complexity being based on the additivity (puzzle-like summation) principle – integration of different parts into a certain whole. The other one consists of the above-described splitting of the perceived object’s image into dozens of quite different object’s features, like shape, movement, colour etc. which are processed separately in distinct cortical areas. The possibility of the coexistence and simultaneous realization of these two antidromic processes within the same anatomic unit (system) is totally incomprehensible, from the classic (neuronal) point of view. On one hand, there is a confluence (merging, junction, maybe synthesis) of a number (hundreds!) of axonal impulses reflecting parts (portions, pieces) of the object, and, on the other hand, disjunction (splitting, breaking, maybe analysis) of the object as a whole – not into parts (portions, pieces) – but into so drastically different (somewhat category-like) and causally disconnected features, like color, form, texture, movement, spatial relationships, etc. And after (or in parallel to) that, there is binding of such disjoined features into an integral coherent image.

Such antidromic way of stimuli processing within the neural networks display the bias/variance dilemma that is incompatible with the neuronal theory, that presenting an insuperable obstacle for the theory of computational mechanism of the brain functioning. According to the modern computational language, the problem is formulated as claiming that the objects or their different aspects have to be represented (to co-exist) within the same physical “hardware” (brain), that resulting in the “superposition catastrophe”\textsuperscript{20}. Therefore, if to define the main postulate of the neuronal theory in the way that each mental representation (“symbol”) of the external objects is represented by the corresponding subsets of coactive neurons within the same brain structure, then if more than one of such “symbols” become active at a moment (which must occur within the real functioning pattern), they become superimposed by co-activation (structural “overlapping” of the activated subsets). In such a case, any information carried by the “overlapping” subsets must be lost, that being the mortal verdict for the whole foundations of the classic neuronal theory.

Consequently, the binding problem, being not a purely theoretical construction but arisen from the very heart of the neurobiological and psychological reality, looks incomprehensible within the framework of the anatomic structure and physiological regularities associated with the perception process. An especial question concerns the relation of the binding problem per se (apart from the above integration of signals within the retina’s neurons) to the somatic-mental gap, which seems so fathomless. In this respect, the main question concerns the ‘localization’ of the events associated with splitting of the signals into different components (constituents, features) and their consequent binding, or, better to say, timing of this signal processing in relation to the somatic-mental gap (before or after?). A reasonable supposition is that, the integration of the initial signals within the retina level precedes their further splitting which, hence, occurs still within the somatic level (thalamo-cortical framework), i.e. before the gap, while the binding is realized within the mental level, i.e. after jumping over the gap. The in principle possibility that both the splitting and binding take place within the mental sphere (after the gap) is hardly probable since the anatomical areas in the brain cortex correspond to the already split components/features\textsuperscript{11-17} so that the splitting seems to occur before the signals reach the final 5\textsuperscript{th} (cortical) neuron.
2. The "hard problem" of consciousness

Recently, the psycho-physical gap was expressed in the form of "hard problem" of consciousness\(^\text{21,22}\) that exposes a naively simple question: how and why performance of any form of neural activity can give rise to subjective experience. Accordingly, the "hard problem" has centralized two main questions: (i) how at all consciousness can emerge from a physical system of whatever degree of complexity; and (ii) how the emerged consciousness can make any effect upon physical stuff via efferent neural ways. By critical consideration of a number of works on consciousness of theoretical significance\(^\text{23-34}\), Chalmers demonstrates their insolvency concerning the "hard problem"\(^\text{21}\). The question has a more general meaning: why should a physical system of any degree of complexity give rise to experience at all?

In the detailed description of the "hard problem", Chalmers\(^\text{21}\) particularizes the 'why' and 'how' questions which may look like emotional variations expressing different aspects of the same main question, e.g. "why should physical processing give rise to a rich inner life at all?"; "why doesn't all this information-processing go on 'in dark'?"; "why is it that when our cognitive systems engage in visual and auditory information-processing, we have visual or auditory experience...?"; "why is it that when electromagnetic waveforms impinge on a retina"... the "discrimination and categorization is experienced as a sensation of vivid red!?"; "why is the performance of these functions accompanied by experience!?"; "...you have explained how information is discriminated ... but you have not explained how it is experienced"; "how can we explain why there is something it is like to entertain a mental image, or to experience an emotion?"; and finally: "It is widely agreed that experience arises from a physical basis, but we have no good explanation of why and how it so arises" (\textit{boldface} is mine). In the above utterances, \textit{why} and \textit{how} (not much differing from each other) are not specifically loaded with a strict terminological burden but eloquently express "what the problem is in the first place"\(^\text{35}\). In this respect, Chalmers' \textit{why} and \textit{how} questions diverge from particularly classified by Guzeldere\(^\text{36}\) four \textit{w-questions (what, where, who, and why)} and the further-how question. In the latter classification, all the \textit{w-questions} (including \textit{why}) correspond to the Chalmers' 'easy problems' while the further-how question corresponds to the 'hard problem'\(^\text{36}\). However, Chalmers' connotation was used in the current debates on the 'hard problem', e.g. 'The why of consciousness' by V. Hardcastle\(^\text{37}\).

The experimental evidence

The experimental approach to the solution of the "hard problem" and the binding problem meets considerable difficulties. The major experimental attempts concern the so-called neural correlates of consciousness. There has recently been a hot dispute about it\(^\text{38}\) including a special issue of the Journal of Consciousness Studies with the target paper by Noë and Thompson\(^\text{39}\). Consequently, the suggested neural correlates of consciousness have included: 40-Hertz oscillations in cerebral cortex\(^\text{29}\), 40-Hertz rhythmic activity in thalamocortical systems\(^\text{40}\), re-entrant loops in thalamocortical systems\(^\text{27}\), extended reticular-thalamic activation system\(^\text{41}\), neural assemblies bound by NMDA\(^\text{*}\)\(^\text{42}\), certain neuron-chemical levels of activation\(^\text{43}\), intralaminar nuclei in thalamus\(^\text{44}\), neurons in the extrastriate visual cortex projecting to prefrontal areas\(^\text{45}\), visual processing within the ventral stream\(^\text{46}\), certain neurons in the inferior temporal cortex\(^\text{47}\), and some other candidates\(^\text{48-51}\). An especial topic concerns neuronal intracellular microtubules’ network as one of the candidates for being a correlate of consciousness\(^\text{52-60}\).

However, to my mind, the major contribution has been made by E. R. John and his school\(^\text{61-69}\) who clearly formulated some particular actual topics to be studied in the sphere of the experimental neuroscience. The questions included the \textit{synchronization} of activities of different neural units within a brain area, coherent activity between the brain areas, the role of coherent oscillations in binding, the functional significance of distributed coherence, spatiotemporal patterns of coherence, and the role of coherence in brain encoding of information\(^\text{62}\). John’s substantial contribution to the binding problem investigation is grounded on solid, consistently accumulated experimental evidence obtained on animals and humans (cognition, psychiatric disorders, and anesthesiology) grounded on the modern sophisticated technology and brilliant neurosurgical mastery and appearing in consistent accordance with the results of other teams worldwide.

The spatiotemporal arrangement (coherence) in brain encoding of information\(^\text{63}\) is of an especial...*The NMDA is a receptor for the neurotransmitter glutamate, which is the most important excitatory transmitter in the brain.*
interest being in connection with my own studies on memorization process\textsuperscript{70}. This concerns, in particular, a certain time incompatibility: “[F]ading but persistent recollection of the recent past coexists in subjective continuity with the momentary present”\textsuperscript{63}. Such conclusion has necessitated suggestion about constant evaluation of the current sensory information immediately confronted with the previous perceptual frame (emotional/motivational relevance of the episodic memories, the whole background etc.). However, any present immediately (instantly!) becomes the past, so that this past is continually including and accumulating all the current Present. However, each present’s moment contains sensory packet which, while instantly joining the past, immediately turns into perceptual packet.

The experimentally demonstrated synchronization of neuronal activity within and coherence among spatially separated brain regions was supposed to be a factor binding distinct attributes into an integral percept\textsuperscript{71-73}. Such binding was considered as a prerequisite for the subjective awareness\textsuperscript{62,74}. In this respect, the facts of importance are those connected with oscillations synchronized across the brain regions that have been found in human intracranial recordings enhanced during cognitive tasks and abolished by anesthesia\textsuperscript{75,76}. The essential finding is that these phase-locked oscillations occur, as it has been emphasized, with zero time delay between the involved regions\textsuperscript{77}.

Thus, such a synchrony between distant brain regions with zero time delay is unexplainable by discrete synaptic transactions, which require appreciable time, while the field influence spreads instantly. This is the ground for integration of the respective different simultaneous and consecutive neural processes covering a certain whole that is emphasized by John as a crucial theoretical obstacle for any comprehensive theory of consciousness: “[I]ntegration of this dispersed information into global consciousness is not compatible with the capacities of any single cell”\textsuperscript{62}. Together with this, John does not accept a proposition about the existence of a remote common source from which “parallel influences” may arrive simultaneously at all the areas involved in analysis of the attributes of the corresponding signal. He claims that such “hypothetical generator must be credited with a priori knowledge of what features are to be bound, requiring multimodal sensitivity to those attributes which are fractionated. This paradox rules out a common source as a plausible explanation…”\textsuperscript{62}. Such logical conclusion was supported by the experiments using multiple moving microelectrodes chronically implanted into conditioned cats\textsuperscript{78}. It has been shown that the observed covariance was due to local neuronal activity rather than to volume conduction from some common source in a distant region.

Field principle and extra ingredient

The field concept looks as a highly promising notion for explanation of consciousness. The modern theories of consciousness using the field principle fall into two groups: (i) those based on the established physical fields, and (ii) those grounded on autonomous fields irreducible to the established physical fundamentals. The theories of the 1st group are based on the reductionist approach explaining consciousness from the physicalist point of view; they may, in principle, elucidate only physical basis of the neural correlates of consciousness\textsuperscript{35,51,79}, leaving the ‘hard problem’ unsolved. The theories of this group are based mainly on the electromagnetic field\textsuperscript{62,63,80-88} and quantum mechanical approach\textsuperscript{52-60} implicating the field principle. In particular, the electromagnetic field conception elaborated by John explains the peculiarities of the binding problem by analyzing and concretizing the stimuli processing (fractionation, selection, and reassembly of the signals).

As to the theories based on the autonomous field, i.e. irreducible to the established physical fundamentals, they look as a promising tool for solving the ‘hard problem’. However, in the existing field theories, namely, “morphic resonance” by R. Sheldrake\textsuperscript{89,90}, “integrated field” by M. Kinsbourne\textsuperscript{91}, “mental field” by B. Libet\textsuperscript{92}, “receptive field” by G. Rosenberg\textsuperscript{93}, “field of consciousness” by W. Hasker\textsuperscript{94}, and “unified conscious field” by J. Searle\textsuperscript{87,88}, the field notion, as it has been comprehensively described previously\textsuperscript{95}, is not specified and has only symbolic meaning. All those theories, although different in content, have a common denominator: the notion of field in each theory has no field-specific postulations.

Thus, the concluding Chalmers’ claim that “a full theory of consciousness must build an explanatory bridge” between neurophysiology and consciousness\textsuperscript{21} could be achieved if an extra ingredient in the explanation is included, because “the emergence of experience goes beyond what can be derived from the physical theory”\textsuperscript{21}. Consequently, the concept of the
extra ingredient, itself, means that the consciousness is irreducible to the established physical fundamentals. On the basis of these clearly dualistic utterances**, Chalmers dwells on possible candidates for such extra ingredient, namely, non-algorithmic processing, non-linear and chaotic dynamics21, quantum mechanics52, and even “future discoveries” in neurophysiology. He convincingly proves that there is “nothing extra” in those ingredients as far as their explaining power is concerned – still the same question remains unanswered: why should each of those processes give rise to the experience?

5. Discussion
The above considerations lead to the following conclusions.

1. The results by Desmedt and Tomberg77 showing zero time delay in phase locked oscillations between the involved brain regions are of especial importance explaining the existing synchrony of neural processes between distant brain regions. Such a zero time delay is unexplainable by discrete synaptic transactions and can be explained by accepting the field principle characterized by the instant speed (speed of light).

2. However, the electromagnetic theories of consciousness have to be accepted as groundless that is proven even by S. Pockett82 – one of the theory initiators80,81. Moreover, in general, the idea of the electromagnetic "nature" of consciousness stands in contrast to biological reality.

3. In the field theories of consciousness based on the irreducible fields, the notion of “field” becomes to be either tautological, or metaphoric, or esoteric, or allegoric, or symbolic, so that in the expression "field of consciousness", the word “field” loses its scientific meaning and could be easily exchanged with the words “state”, “stream”, “process”, “phenomenon”, etc.

4. However, in spite of the unfitness of the idea about the physical (electromagnetic) nature of consciousness as well as the inconsistency of the suggested irreducible field-based theories of consciousness, the field notion, which is the only principle explaining the non-locality, wholeness, and synchronicity of the conscious events, must be accepted at least as an abstract principle with the hope for its possible realization.

5. The only theoretical possibility concerns the irreducible field-based theory of consciousness, in which the field notion is strictly defined by the formal prerogative field-specific postulates inherent to any "non-metaphorical" field, irrespective to its physical nature. The formulated postulates98-100 include implication of actual field source(s); action-at-a-distance; unboundedness (absence of any distinct boundaries of the field influence, which, hence, theoretically is infinite); field integrity; dependence of certain parameters of any in-field-occurring process on its coordinates within a certain whole; field ‘direction’ from the ‘classic’ scalar and vectorial fields till the modern highly complicated (tensor-string-twistor, etc.) expressions associated with the ‘over-classic’ number of dimensions; and field measurability, i.e. quantitative estimation of certain field particulars, like intensity, decrement etc. The acceptance of the above postulates secures any suggested field conception from sliding into symbolic-metaphoric-allegoric meaning.

6. Consequently, the problem consists in formulation of a non-physical extra ingredient described by the strict physical-mathematical language and associated with the biological reality. The corresponding irreversible field-based theory of consciousness must be grounded on the above-formulated postulates. At the first glance, the task seems utopian.

7. However, a miracle is that exactly such field theory was elaborated rather long ago101-103 comprehensively reviewed104-107, and employed for the explanation of the consciousness enigma98-100.

References
5 Toward a science of consciousness (Consciousness and its place in nature), Conference Research Abstracts in J Cons Stud, Skövde, Sweden, August 7-11, 2001, 1.
6 Toward a science of consciousness, Conference Research Abstracts in the J Cons Stud, Tucson, USA, April 8-12, 2002, 1.

**”This position qualifies as a variety of dualism, as it postulates basic properties over and above the properties invoked by physics. But it is an innocent version of dualism, entirely compatible with the scientific view of the world. Nothing in this approach contradicts anything in physical theory; we simply need to add further bridging principles to explain how experience arises from physical processes. … If the position is to have a name, a good choice might be naturalistic dualism”***
7 Toward a science of consciousness, Conference Research Abstracts in the J Cons Stud, Tucson, USA, April 7-11, 2004, 1.
8 Toward a science of consciousness, Conference Research Abstracts in the J Cons Stud, Tucson, USA, April 4-8, 2006, 1.
18 Damasio A R, The brain binds entities and events by multiregional activation from convergence zones, Neural Comp, 1 (1989) 123.
51 Crick F & Koch C, Feature article: Consciousness and neuroscience, Cerebral Cortex, 8 (1998) 97.
53 Hameroff S, Fundamental geometry: The Penrose-Hameroff Orch OR model of consciousness, in Geometry and the

107


55 Hameroff S., Consciousness, the brain, and space-time geometry, in Cajal and consciousness: Scientific approaches to consciousness on the centennial of rámón y cajal's textura, edited by P C Marijuan, Ann NY Acad Sci, 929 (2001) 74.


59 Penrose R, Consciousness, the brain, and space-time geometry: an addendum; Some new developments on the Orch OR model for consciousness, in Cajal and consciousness: scientific approaches to consciousness on the centennial of rámón y cajal's textura, edited by P C. Marijuan, Ann NY Acad Sci, 929 (2001) 74.


61 John E R, Machinery of the mind (Birkhauser, Boston, MA) 1990.


64 John E R, Easton P & Isenhart R, Consciousness and cognition may be mediated by multiple independent coherent ensembles, Cons & Cognit, 6 (1997) 3.


