

# Advances in Interdisciplinary Researches to Construct a Theory of Consciousness

Pierre R. Blanquet

*Club d'Histoire des Neurosciences, University P. & M. Curie, Paris, France*

*E-mail: [pr.blanquet@free.fr](mailto:pr.blanquet@free.fr)*

*Received August 22, 2011; revised October 13, 2011; accepted October 28, 2011*

## Abstract

The interdisciplinary researches for a scientific explanation of consciousness constitute one of the most exciting challenges of contemporary science. However, although considerable progress has been made in the neurophysiology of states of consciousness such as sleep/waking cycles, investigation of subjective and objective nature of consciousness contents raises still serious difficulties. Based on a wide range of analysis and experimental studies, approaches to modeling consciousness actually focus on both philosophical, non-neural and neural approaches. Philosophical and non-neural approaches include the naturalistic dualism model of Chalmers, the multiple draft cognitive model of Dennett, the phenomenological theory of Varela and Maturana, and the physics-based hypothesis of Hameroff and Penrose. The neurobiological approaches include the neurodynamical model of Freeman, the visual system-based theories of Lamme, Zeki, Milner and Goodale, the phenomenal/access hypothesis of Block, the emotive somatosensory theory of Damasio, the synchronized cortical model of Llinas and of Crick and Koch, and the global neurophysiological brain model of Changeux and Edelman. There have been also many efforts in recent years to study the artificial intelligence systems such as neurorobots and some supercomputer programs, based on laws of computational machines and on laws of processing capabilities of biological systems. This approach has proven to be a fertile physical enterprise to check some hypothesis about the functioning of brain architecture. Until now, however, no machine has been capable of reproducing an artificial consciousness.

**Keywords:** Models of Consciousness, Third and First-Person Data/Qualia, Neuronal Workspace, Thalamo-Cortical Neurons, Artificial Intelligence

## 1. Introduction

By separating the immaterial mind and the material body, Descartes (1596-1650) traced a dividing line between two incommensurable worlds: the spiritual and the material world [1]. From this ontological dualism is born an epistemological dualism according to which the material must be known by science and the mind by introspection. During the 18<sup>th</sup> century, the definition of the mind started to emerge with that of consciousness, recognized as instrument of knowledge of the world which surrounds us as also of our interiority. Hume (1711-1776), for example, defined the mind: "as an interior theatre on which all that is 'mental' unravels in a chaotic way in front of an interior eye whose eyelids would not blink" [2]. On the contrary, the materialist philosophers of the 18<sup>th</sup> century will give up the cartesian immaterial substance, but they will keep the metaphor of the body-machine and will extend this

metaphor to all human functions, including thought. Far from showing the immateriality of the heart, indeed, the cartesian cogito would prove rather for La Mettrie (1709-1751), Diderot (1713-1784) and Holbach (1723-1789) that the matter can think. In this line of idea, transformist/proto-evolutionist and darwinian theories of 18<sup>th</sup> and 19<sup>th</sup> centuries, by introducing the idea of a biological origin of the man, introduced the vast field of investigation of the "living matter" and the problem of "when" and "how" of the mind [3].

However, it was necessary to await the arise of the psychology in the mid-nineteenth century, so that consciousness becomes the central object of a new discipline claiming science and asserting its independence with regard to philosophy. This introspective approach of mind dominated the investigations of researchers such as Herman von Helmholtz (1821-1894), Wilhelm Wundt (1832-1920) and William James (1842-1910) [4-6]. In the end of the

19<sup>th</sup> century, in his founder text for a phenomenology, Husserl (1859-1938) exposed his thesis on the nature of consciousness [7]. Following Brentano's work, he adopted the concept of intentionality and gave it a central role concerning the transcendental ego. He taught that consciousness is never empty: it always aims an object and, more generally, the world. The ego always carries in it the relationship with the world as intensional aiming. The early 20<sup>th</sup> century saw the eclipse of consciousness from scientific psychology [8,9]. Without denying the reality of subjective experiences, the strategy of behaviorists was to draw aside consciousness of the direct field of investigation and to regard the brain as a "puzzle-box". They then put "between brackets" the irresolute problems which encumbered the psychology to study stimuli and answers, *i.e.* behaviors which reduce the analysis to rigid relations between inputs (stimuli) and outputs (movements). In the 1960s, the grip of behaviorism weakened with the rise of the cognitive psychology [10,11]. However, despite this renewed emphasis on explaining cognitive capacities such as memory, perception and language comprehension, consciousness remained a largely neglected topic until the major scientific and philosophical researches in the 1980s. Researches on the correlations between the identifiable mental activities and the objectivable cerebral activities succeeded the traditional questioning on the relation between the mind and the brain. They were approached by neurophilosophers, neurobiologists and researchers in artificial intelligence as familiar with the epidemiologic and ontological questions as with the methodologic and empiric ones. My purpose here is to review the main attempts to provide an adequate explanatory basis for consciousness.

## 2. Consciousness, a Challenge for Philosophy and Science

### 2.1. Difficulties in Philosophically Tackling the Problem of Consciousness

Traditionally, one supposed in philosophy of mind that there was a fundamental distinction between the dualist philosophers, for which there are two kinds of phenomena in the world, the mind and the body, and the monist philosophers, for which the world is made of only one substance. Although dualism can be of substance, as Descartes thought it, the majority of dualist philosophers currently adopt a dualism of property which admits that the matter can have both mental and physical properties. Likewise, if the monism can be an idealism (in the sense of the philosopher Berkeley), the totality of monist philosophers in the present time are materialists. Dualism of property asserts the existence of conscious properties that are neither identical with nor reducible to physical pro-

erties, but which may nevertheless be instantiated by the very same things that instantiate physical properties. The most radical dualism is the "psychological parallelism", which seeks to account for the psychophysiological correlations while postulating that the mental state and the cerebral state correspond mutually without acting one on the other [12]. A less radical version, "the epiphenomenalism", recognizes the existence of causal influences of brain on the mental states, but not the reverse [13]. In contrast to dualism, the monism claims that everything that exists must be made of matter. A modern version of materialism is the physicalism. One type of physicalist materialism is the radical current of thought called "eliminativism", which rejects the very notion of consciousness as muddled or wrong headed. It affirms that the mental states are only temporary beliefs intended to be replaced by neurobiologic models of research [14]. Another form of physicalist materialism is the thesis of the strong identity. It affirms the existence of an internal principle of control (the mental principle) which is anything else only the brain: the mental is reducible with the biological properties, which themselves are reducible with physics. However, the thesis most commonly adopted is that of functionalism. The functionalism has been proposed to answer an intriguing question: how to explain, if the theory of identity is admitted, that two individuals can have different cerebral states and nevertheless to have at one precise time exactly the same mental state? The response of functionalists is as follows: what there is identical in the two different cerebral occurrences from the same mental state, it is the function. Whatever the physical constitution of the cerebral state, several mental states are identical if their relations are causal. For the functionalists, the mental states are thus defined by their functional role inside the mental economy of the subject [15].

To accept traditional dualism it is agree to make a strict distinction between the mental and the physical properties. In other words, it is to give up the unified neuroscientific theory which one can hope for to obtain one day. But to accept the solutions recommended by the physicalist materialism is worse still, because those end up denying the obvious fact that consciousness has internal subjective and qualitative states. For this reason, certain philosophers tried to solve the problem, either by adopting a mixed theory, both materialistic and dualistic, or by denying the existence of subjective states of consciousness. David Chalmers, for example, adopted the first solution. Chalmers is known for his formulation of the distinction between the "easy" problem of consciousness and the single "hard" problem [16,17]. The essential difference between the easy problem and the hard (phenomenal) problem is that the former is at least theoretically answerable via the standard strategy of functionalism. In

support of this, in a thought experiment, Chalmers proposed that if zombies are conceivably complete physical duplicates of human beings, lacking only qualitative experience, then they must be logically possible and subjective personal experiences are not fully explained by physical properties alone. Instead, he argued that consciousness is a fundamental property ontologically autonomous of any known physical properties. Chalmers described therefore the mind as having “phenomenal” and “psychological” aspects. He proposed that mind can be explained by a form of “naturalistic dualism”. That is, Chalmers accepted the analysis of functionalists while introducing the concept of consciousness irreducible into his system. According to him, since the functional organization provides the elements of mental states in their nonconscious forms, it is necessary that consciousness is added again to this organization.

Other philosophers, on the contrary, claimed that closing the explanatory gap and fully accounting for subjective personal experiences is not merely hard but rather impossible. This position was most closely associated with that, for example, of Colin McGinn, Thomas Nagel [18,19] and Daniel Dennett [20]. In particular, Dennett argued that the concept of qualia is so confused that it cannot be put to any use or understood in any non-contradictory way. Having related the consciousness to properties, he then declared that these properties are actually judgements of properties. That is, he considered judgements of properties of consciousness to be identical to the properties themselves. Having identified “properties” with judgements of properties, he could then show that since the judgements are insubstantial, the properties are insubstantial and thence the qualia are insubstantial. Dennett concluded therefore that qualia can be rejected as non-existent [20]. For Dennett, consciousness is a mystifying word because it supposes the existence of a unified center of piloting of thoughts and behaviors. For him, psychism is a heterogeneous unit which combines a series of mental processes that one knows little about such as perception, production of the language, training, etc...: we allot, with the others and ourself, some “intentions”, a “consciousness”, because our behaviors are finalized [21,22]. In other writings, Dennett defended an original thesis on the free will [23]. For that, he rested on evolutionary biology, on cognitive sciences, on the theory of cellular automata and on memetics [24,25]. Taking the opposite course to the argument of those which say that evolutionary psychology, together with memetics, implies necessarily a world deprived of any possibility of choice, Dennett estimated on the contrary that there is something of special in the human subject. According to him, the theory of evolution supports the selective advent of avoiding agents. These agents have the capacity to extract information of the

environment to work out strategies in order to avoid the risks and to choose the judicious behaviors. Now, the performance of this capacity is much more important for the man than for the animal because each human individual memorizes an important quantity of social and cultural informations. It is therefore absurd to conceive a linear determinism similar to that of the “demon” of Laplace. For Dennett, the free will must rather be designed like the chaotic determinism of a vast neuromimetic network, in which informations received in entries are combined according to their respective weights to give completely unforeseeable and nonreproducible, but non random, outputs.

## 2.2. Difficulties in Scientifically Tackling the Problem of Consciousness

One of difficulties in scientifically tackling the problem raised by consciousness comes from the obvious fact that it is the originating principle from which are generated the categories of the interpersonal world and the subjectivity of each personal world. Consciousness is not a substance but an act in itself non objectivable; it escapes any from representation. What consciousness has disconcerting, notes Edelman, it is that it does not seem to raise of the behavior. It is, quite simply, always there [26]. Another difficulty comes owing to the fact that consciousness is a *Janus Bifron*, in the sense that it has both ontology with the first-person perspective and ontology with the third-person perspective, irreducible one with the other [27]. It has a first-person perspective because the mental states are purely subjective interior experiences of each moment of the life of men and animals. Principal sorts of first-person data include visual, perceptual, bodily and emotional experiences, mental imagery and occurrent thoughts. But consciousness has also ontology with the third-person perspective that concerns the behavior and the brain processes of conscious systems [22]. These behavioral and neurophysiological data relevant to the third-person perspective provide the traditional material of interest for cognitive psychology. Principal sorts of third-person data include perceptual discrimination of external stimuli, integration of information to do with the across sensory modalities, automatic and voluntary actions, levels of access to internally represented information, verbal reportability of internal states and differences between sleep and wakefulness. The problem to scientifically approach the consciousness phenomenon is much more complex still because the first-person experience is common for qualia (singular ‘quale’), intentionality and self-consciousness. Introduced by Lewis [28], the term of qualia is either used in contemporary usage to refer to the introspectively accessible phenomenal aspects of our mental live [29], or used in a more restricted way so that qualia are intrinsic

properties of experiences that are ineffable, nonphysical, and given incorrigibly [30], or still used by some philosophers such as Whitehead, which admits that qualia are fundamental components of physical reality and described the ultimate concrete entities in the cosmos as being actual “occasions of experience” [31]. Experiences with qualia are involved, for example, in seeing green, hearing loud trumpets or tasting liquorice, bodily sensations such as feeling a twinge of pain, feeling an itch, feeling hungry, etc... They are also involved in felt reactions, passions or emotions such as feeling delight, lust, fear or love, and felt moods such as feeling elated, depressed, calm, etc... [32,33]. Intentionality and qualia necessarily coexist in the generation of conscious states, but the aspect “qualia” may be distinguished from the aspect “intentionality” insofar as the perception of an object, the evocation of a memory or an abstract thought can be accompanied by different affective experiences (joy or annoys, wellbeing or faintness, etc...) [34].

### 3. Experimental Approaches to a Science of Consciousness

#### 3.1. Neurophysiological Studies of Neural Networks and Neuro-Mental Correlations

There are two common, but quite distinct, usages of the term of consciousness, one revolving around arousal or states of consciousness and another around the contents of consciousness or conscious states. States of consciousness are states of vigilance (*i.e.*, the continuum of states which encompasses wakefulness, sleep, coma, anesthesia, etc) with which attention is associated. They are cyclic and can last several hours. In contrast, conscious states (percept, thought, memory or subjective experiences such as qualia) are not cyclic nor reproducible and can last only a few minutes, seconds or sometimes milliseconds. Conscious states are situated in a spatiotemporal context. They may refer to the past or the future but are always experienced in the present. They introduce the conscious representations and, more or less explicitly, global models of self, alter-ego and world. States of consciousness and contents of consciousness offer therefore very unequal difficulties to the neuroscientific investigation [35-37].

##### 3.1.1. Neurophysiological Studies of States of Consciousness

Considerable progress has been made during the last decade in the neurophysiology of states of consciousness. In particular, impressive progress has been made in the neurophysiological investigation of states of vigilance. Notably, the molecular mechanisms of distinct sleep/wake cycles have been thoroughly studied [38]. The neuroanatomical

systems, the cellular and molecular mechanisms, and the principal types of neurotransmitters involved in these mechanisms for the most part have been identified. In particular, an important line of research has investigated the arousal in altered states of consciousness, for instance, in and after epileptic seizures, after taking psychedelic drugs, during global anesthesia or after severe traumatic brain injury. These studies demonstrate that a plethora of nuclei with distinct chemical signatures in the thalamus, mid-brain and pons, must function for a subject to be in a sufficient state of brain arousal to experience anything at all. In particular, sleep/wake cycles essentially depend on anatomical system which comprises structures of brainstem, thalamic and hypothalamic nuclei, and the nucleus of Meynert. Awakening into the vigilant state correlates with a progressive increase in regional cerebral blood flow, first in the brainstem and thalamus, then in the cortex with a particularly important increase in prefrontal-cingulate activation and functional connectivity [38,39]. Anesthetic, sleep, vegetative state and coma are all associated with modulations of the activity of this thalamocortical network. Although most dreams occur in paradoxical sleep, the neurobiological mechanisms of wakefulness and paradoxical sleep are identical at the thalamocortical level; the only difference between the two states is at the brainstem level. What differentiates these states is the relationship with the environment: wakefulness brings into play motor and sensory interactions with the external world, while the paradoxical sleep is relatively independent of the external world. Thus, deaming may be considered to be an incomplete form of consciousness, uncontrolled by the environment, mainly reflecting internal factors [38].

##### 3.1.2. Neurophysiological Studies of Contents of Consciousness

Contrary to the study of consciousness states, the study of consciousness contents raises very many problems. Much difficulties are due to the brevity and not very reproducible nature of subjective experiences. In addition, the mechanisms of conscious thoughts often result from processes at the same time conscious and unconscious which coexist and even interact (the language, for example, brings into play the unconscious use of linguistic routines). To explain the contents of consciousness of the third-person perspective, we need actually to specify the neural and/or behavioral mechanisms that perform the functions [27]. The availability of behavioral data is reasonably straightforward, because researchers have accumulated a rich body of behavioral data relevant to consciousness. But the body of neural data that has been obtained to date is correspondingly more limited because of technological limitations [33,40]. To study neural mechanisms, researchers use currently a variety of neuroscientific measure-

ment techniques including brain imaging via functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) technology, single-cell recording through insertion of electrodes, and surface recording through electroencephalography (EEG) and magnetoencephalography (MEG) [27,41,42]. However, though these approaches seem quite promising, many experimental findings proved to be not univocal results and must be compared and integrated with findings obtained from other approaches. For example, when one sees a face there be much activities (for example, on the retina and in the early visual cortex) that seem explanatorily redundant for the formation of the conscious percept of the face [37]. To specify precisely the neuronal basis of such a conscious perception, psychologists have perfected a number of techniques (masking, binocular rivalry, continuous flash suppression, motion-induced blindness, change blindness, inattention blindness) [43]. In this design, one keeps as many things as possible constant, including the stimuli, while varying the conscious percept, so that changes in neural activation reflect changes in the conscious percept rather than changes in the stimuli. For instance, a stimulus can be perceptually suppressed for seconds of time: the image is projected into one of the observer's eyes but is invisible, not seen. In this manner, the neural mechanisms that respond to the subjective percept rather than the physical stimulus can be isolated, permitting the footprints of visual consciousness to be tracked in the brain. In some perceptual illusion experiments, on the contrary, the physical stimulus remains fixed while the percept fluctuates. A good example is the Necker cube whose 12 lines can be perceived in one of two different ways in depth.

Contrarily to techniques used for studying the third-person experiences, methodologies for investigating first-person experiences (in particular qualia) are relatively thin on the ground, and formalisms for expressing them are even thinner. The most obvious obstacle to the gathering of first-person data concerns the privacy of such data. Indeed, first-person data are directly available only to the subject having these experiences. To others, these first-person data are only indirectly available, mediated by observation of the subject's behavior or brain processes. In practice, the most common way of gathering data about the conscious experiences of other subjects is to rely on their verbal reports. However, verbal reports have some limits. Some aspects of conscious experience (e.g. the experience of emotion) are very difficult to describe. Moreover, verbal reports cannot be used at all in subjects without language, such as infants and animals. In these cases, one needs to rely on other behavior indicators. For example, if an individual presses one of two buttons depending on which of two ways the creature perceives a Necker cube at a time, this button pressing is

a source of first-person data. A second obstacle is posed by the absence of general formalisms with which first-person data can be expressed. Researchers typically rely either on simple qualitative characterizations of data or on simple parameterization of ones. These formalisms suffice for some purposes, but they are unlikely to suffice for the formulation of systematic theories [27,33].

### 3.1.3. Approaches to Research Animal Consciousness

There is now abundant and increasing behavioral and neurophysiological evidence consistent with, and even suggestive of, conscious states in some animals. This is due to the fact that human studies involving the correlation of accurate reports with neural correlates can provide a valuable benchmark for assessing evidence from studies of animal behavior and neurophysiology of some animals [44]. Relevant forms of report included analysis of responses to binocular rivalry in the case of primates, vocalization in the case of birds such as african grey parrots, and coloration and body patterning in the case of cephalopods such as *Octopus vulgaris*. Rhesus macaque monkeys, for example, were trained to press a lever to report perceived stimuli in a binocular rivalry paradigm. The results from these studies were consistent with evidence from humans subjected to binocular rivalry and magnetoencephalography. They suggested that consciousness of an object in monkeys involves widespread coherent synchronous cortical activity. Likewise, similarities have been found among the functional circuitry underlying the organization and sequencing of motor behaviors related to vocalization in birds and mammals capable of vocal learning. Much of the neural basis for song learning in some birds was found to reside in an anterior forebrain pathway involving the basal ganglia, in particular a striatal neuronal area resembling that present in the mammalian striatum. These homologies were strongly suggestive of neural dynamics that support consciousness in birds. The case of cephalopod mollusks is much less clear. Indeed, cephalopod such as *Octopus* possesses a large population of sensory receptors (they communicate with a nervous system containing between 170 and 500 million cells) and numerous nucleus-like lobes in its brain. Its optic lobe, containing as many as 65 million neurons, plays a critical role in higher motor control and establishment of memory. Moreover, a number of other lobes appear to be functionally equivalent to vertebrate forebrain structures, though their organization bears little resemblance to the laminar sheets of mammalian cortex. Recently, laboratory tests and observations in a natural environment showed that *Octopus* is heavily dependent on learning, and even might form simple concepts. This strongly suggested that cephalopod mollusks have a certain form of primary consciousness [45].

### 3.2. Approaches to Build Artificial Intelligence Systems

Artificial Intelligence has not been sparing of metaphors concerning the functioning of the human mind [46]. In 1950, Alan Turing tackled the problem of computationalism by proposing its famous test to establish whether one can consider a machine as intelligent as a human [47]. So far, however, no computer has given responses totally indistinguishable from the human responses. It appears that this computational cognitivism is limited insofar as it is founded on the formal character of calculation (in this case, to think is equivalent to process data, *i.e.* to calculate, to handle symbols). This neuroscientific approach is not very different from that where computer simulations and mathematical models are used to study systems such as stomachs, planetary movements, tornadoes, and so on. In contrast, connectionism equates thinking with the operation of a network of neurons and argues that every cognitive operation is the result of countless interconnected units interacting among themselves, with no central control. The connexionist networks are in general adaptive and allow the study of mechanisms of training [48,49]. Such an approach is based on the view that the nervous system itself computes [50]. In this topic, new research has particularly developed that consists in making work equations put in darwinian competition thanks to evolutionary algorithms inspired by the modeling of certain natural systems (for example, the competition between the social insects in the construction of the nest) [51-53]. These systems, which are generic population-based metaheuristic optimization algorithms, are able to solve problems using some mechanisms inspired by the biological evolution such as the reproduction, mutation, recombination and selection. The principle on which they are founded consists in initializing a population of individuals in a way dependent on the problem to be solved (environment), then to make move this population of generation in generation using operators of selection, recombination and change. Actually, a number of cyberneticians try to approach this difficult problem with systems multi-agents (SMA), or "Massively Multi-Agents" [54-60]. An agent-based model is a class of computational models for simulating the actions and interactions of autonomous agents (both individuals or collective entities such as organizations or groups) with the view to assessing their effects on the system as a whole. This model simulates the simultaneous operations and interactions of multiple agents located in an environment made up of objects which are not agents and unlikely of evolution, in an attempt to re-create and predict the appearance of complex phenomena. The agents can thus substitute for the programmer and even produce unexpected results. In

this vein of approaches, a novel area of great interest is the construction of robotic organisms. One essential property of neurobots is that, like living organisms, they must organize the unlabeled signals they receive from the environment into categories. This organization of signals, which in general depends on a combination of sensory modalities (e.g. vision, sound, taste, or touch), is a perceptual categorization, as in living organisms, makes object recognition possible based on experience but without a priori knowledge or instruction. Like the brain, neurobots operate according to selectional principles: they form categorical memory, associate categories with innate values, and adapt to the environment [61].

An important problem arises however with these new perspectives of research: any constructive and exhaustive approach of artificial consciousness must define a system which has access, as the human primary consciousness, within the meaning of its own knowledge. To investigate this question, as Owen Holland said it, the system must be able simultaneously to produce an intentional representation and the perception in its own organization of this intentionally generated form; it must be self-aware [62]. A question then becomes: when will a machine become self-aware? Although an answer to this question is hazardous, one can determine at least a plausible necessary precondition without which a machine could not develop self-awareness. This precondition derives from the assertion that, to develop self-awareness, a neural network must be at least as complex as the human brain. Recently, an enormous project (Human Brain Project) was given for objective, believing this could be achievable in as little as 10 years time, to succeed in simulating the functioning of the neocortex of mammals by means of the fastest supercomputer architecture in the world, the IBM's Blue Gene platform [63]. For the moment, one modelled a single cortical column consisting of approximately 60,000 neurons and 5 km of interconnecting synapses, chosen from about 15,000 experiments carried out in laboratory. This Blue Brain project requires to reproduce in the future the equivalent of the million functional units that holds the neocortex. It should be however noted that this project refers only to a necessary but not sufficient condition for the development of an artificial consciousness. Even if the machine becomes as skilled as humans in many disciplines, one cannot assume that it has become self-aware. The existence of a powerful computer equipped with millions of gigabytes is not in itself sufficient to guarantee that the machine will be a self-awareness intelligence. Moreover, it remains to define the criteria which will make possible to recognize that an artificial entity has a conscious state. Indeed, the problem is not to know if a machine suffers, but if it behaves "as if" it suffered [64,65].

## 4. Explanatory Theories of Consciousness

Explanatory theories of consciousness should be distinguished from the experimental approaches to phenomena of consciousness. While the identification of correlations between aspects of brain activity and aspects of consciousness may constrain the specification of neurobiologically plausible models, such correlations do not by themselves provide explanatory links between neural activity and consciousness. Models of consciousness are valuable precisely to the extent that they propose such explanatory links. To days, one can arbitrarily classify the various approaches to modeling consciousness into two categories: the theories making correspond certain functional modes of the brain to the conscious activity, and the theories binding the structure of neural networks to conscious activity .

### 4.1. Theories Making Correspond Certain Modes of the Brain to the Conscious Activity

In this category, one finds mainly the phenomenological approaches of Francisco Varela and Humberto Maturana, the physics-based hypothesis of Stuart Hameroff and Roger Penrose, and the neurodynamical model of Walter J. Freeman.

**Model of Varela and Maturana.** Varela, and his mentor Humberto Maturana, developed a model based on the notion of autopoiesis [66-70]. Based on cellular life, autopoiesis attempts to define, beyond the diversity of all living organisms, a common denominator that allows for the discrimination of the living from the non-living. Inside the boundary of a cell, many reactions and many chemical transformations occur but, despite all these processes, the cell always remains itself and maintains its own identity. The consequence is that the interaction between a living autopoietic unit and a component of its environment is only dictated by the way in which this component is 'seen' by the living unit. To characterize this very particular nature of interaction, Maturana and Varela used the term of "cognition". In their theory, the environment has its own structural dynamics but it does not determine the changes in organism. Although it can induce a reaction in the organism, the accepted changes are determined by the internal structure of the organism itself. The consequence is that the environment brings to life the organism and the organism creates the environment with its own perceptory sensorium. It should be emphasize that this thinking is close to certain european philosophies, in particular to that of Merleau-Ponty [71]. To express this process of mutual calling into existence, this co-emergence, this equivalence between life and cognition, Varela and Maturana used the word of "enaction". For Varela, the mind as a phenomenology in action, viewed from either

the first- or the third-person perspective, can only be described as a behavior literally situated in a specific cycle of operation [72]. For him, the mind is not in the head, it is in a non-place of the co-determination of inner and outer [73]. There is no separation between the cognitive act and the organic structure of life, they are one ; the traditional cartesian division between matter and mind disappears at the level of human cognition at which the notion of consciousness appears. To signify that human consciousness has its counterpart in the organic structure, that there is no consciousness outside the reality of bodily experience, Varela used the term of "embodiment" [74]. We are therefore global dynamic processes, in dynamical equilibrium, emerging and acting from interactions of constituents and interactions of interactions. In this thesis, the brain level is only considered to contribute to properties of conscious experiences. At the phenomenal level, the constitution of conscious moments implies a high temporal integration of multiple contents emerging in a transitory way. According to Varela, because of its biophysical organization (its organizational closure), the brain belongs to multistable dynamical systems, in which eigenbehaviors are constrained by a landscape of multiple non-stable attractors. There are however some methodological problems in this theory. The first is the old problem that the mere act of attention to one's experience transforms that experience. This is not too much of a problem at the start of investigation, because one has a long way to go until this degree of subtlety even comes into play, but it may eventually lead to deep paradoxes of observership. The second problem is that of developing a language or a formalism in which phenomenological data can be expressed. Indeed, the notorious ineffability of conscious experience plays a role here, because the language one has for describing experiences is largely derivative on the language one has for describing the external world. The third difficulty lies in the failure, or at least the limitations, of incorrigibility: judgments could be wrong.

**Model of Hameroff and Penrose.** For Stuart Hameroff and Roger Penrose, neurons belong to world of traditional physics, are calculable and cannot explain as a such consciousness. They proposed therefore a new physics of objective reduction which appeals to a form of quantum gravity to provide an useful description of fundamental processes at the quantum/classical borderline [75-78]. Within this scheme, consciousness occurs if an appropriately organized system is able to develop and maintain quantum coherent superposition until a specific "objective" criterion (a threshold related to quantum gravity) is reached; the coherent system then self-reduces (objective reduction). This type of objective self-collapse introduces non-computability, an essential feature of consciousness which distinguishes our mind from classical computers.

Objective reduction is taken as an instantaneous event (the climax of a self-organizing process in fundamental space-time) [31]. In this model, quantum-superposed states develop in microtubules subunit proteins (tubulins) within the brain neurons. They recruit more superposed tubulins until a mass-time-energy threshold (related to quantum gravity) is reached. At this point, self-collapse, or objective reduction, abruptly occurs. One equates the pre-reduction, coherent superposition phase with pre-conscious processes, and each instantaneous (and non-computable) objective reduction, or self-collapse, with a discrete conscious event. Sequences of objective reduction give rise to a “stream” of consciousness. Microtubule-associated-proteins can “tune” the quantum oscillations of the coherent superposed states. The objective reduction is thus self-organized, or “orchestrated”. Each orchestrated objective reduction event selects (non-computably) microtubule subunit states which regulate synaptic/neural functions using classical signaling. The quantum gravity threshold for self-collapse is relevant to consciousness because macroscopic superposed quantum states each have their own superposed space-time geometries. However, when these geometries are sufficiently separated, their superposition becomes significantly unstable and reduce to a single universe state. Quantum gravity determines thus non-computably the limits of the instability. In sum, each orchestrated objective reduction event is a self-selection of space-time geometry coupled to the brain through microtubules and other molecules. This orchestrated objective reduction provides us with a completely new and uniquely promising perspective on the hard problem of consciousness [77,78]. The model of Hameroff and Penrose has received serious criticism, notably from certain philosophers such as Rick Grush and Patricia Churchland [79]. These authors pointed out that microtubules are found in all plant and animal cells, and not only in brain neurons. They also stated that some chemicals that are known to destroy microtubules do not seem to have any effects on consciousness and that, conversely, anaesthetics act without affecting the microtubules. Another objection addresses one of the strengths of Penrose and Hameroff’s model, which is, according to its authors, that it can account for the unity of consciousness. Indeed, if this impression of human consciousness unity should prove to be an illusion, then explanations based on non-locality and quantum coherence would become irrelevant.

**Model of Freeman.** The work of Walter J. Freeman was based mainly on electrophysiological recording of the olfactory system of awake and behaving rabbits [80-84]. Freeman found that the central code for olfaction is spatial. Although this had been predicted by others on the basis of studies in the hedgehog, certain aspects of his results were surprising. For example, Freeman showed

that the information is uniformly distributed over the entire olfactory bulb for every odorant. By inference every neuron participates in every discrimination, even if and perhaps especially if it does not fire, because a spatial pattern requires both black and white. He discovered that the bulbar information does not relate to the stimulus directly but instead to the meaning of the stimulus. This means that the brain does not process information in the commonly used sense of the word. When one scans a photograph or an abstract, one takes in its import, not its number of pixels or bits. The brain processes meaning as in this example. He also found that the carrier wave is aperiodic: it does not show oscillations at single frequencies, but instead has wave forms that are erratic and unpredictable, irrespective of odorant condition. In the theory of Freeman, the chaotic activity distributed among the neuronal populations is the carrier of a spatial pattern of amplitude modulation that can be described by the local heights of the recorded waveform. When an input reaches the mixed population, an increase in the nonlinear feedback gain will produce a given amplitude-modulation pattern. The emergence of this pattern is considered to be the first step in perception: meaning is embodied in these amplitude-modulation patterns of neural activity, whose structure is dependent on synaptic changes due to previous experience. Through the divergence and convergence of neural activity onto the entorhinal cortex, the pulse patterns coming from the bulb are smoothed, thereby enhancing the macroscopic amplitude-modulation pattern, while attenuating the sensory-driven microscopic activity. Thus, what the cortex “sees” is a construction made by the bulb, not a mapping of the stimulus. Hence, perception is an essential active process, closer to hypothesis testing than to passive recovery of incoming information. The stimulus then confirms or disconfirms the hypothesis, through state transitions that generate the amplitude-modulation patterns. Finally, through multiple feedback loops, global amplitude-modulation patterns of chaotic activity emerge throughout the entire hemisphere directing its subsequent activity. These loops comprise feedforward flow from the sensory systems to the entorhinal cortex and the motor systems, and feedback flow from the motor systems to the entorhinal cortex and from the entorhinal cortex to the sensory systems. Such global brain states emerge, persist for a small fraction of a second, then disappear and are replaced by other states. It is this level of emergent and global cooperative activity that is crucial for consciousness. Freeman tackled also the enigmatic problem of the nature of the free will. He proposed that the man, entirely engaged in a project and in permanent relation with the other men and the world, takes his decisions in real time by implying all his body. The perceived will as conscious is informed only with a



light delay. Not deciding a behavior in progress, it only takes action to modulate the various aspects of the voluntary decision and to legitimate it taking into consideration whole of significances which constitute the subject. There are however some problems in this interesting Freeman's account [85]. The mechanisms of origine and the control of gamma oscillations are not yet entirely clear. As predicted by Freeman, during gamma oscillations an average lead/lag relation exists between local excitatory and inhibitory cells. Now, recent analyses of the cellular dynamics concluded that recurrent inhibition from fast spiking inhibitory cells is largely responsible for maintaining the rhythmic drive, but the role played by excitatory processes in modulating or driving the oscillations remains undetermined. Since cortical networks form a dilute and massively interconnected network, a satisfactory explanation for gamma activity should explain the onset and offset of gamma activity in relation to events at more distant sites and at larger scale in the brain. Without clarification of these mechanisms it remains difficult to define the link of gamma activity to storage, retrieval and transmission of information, and between thermodynamic and cognitive or informational perspectives.

## 4.2. Theories Binding the Structure of Neural Pathways to Conscious Activity

In this second category of models, a first strategy consists in being focused on the visual system. This approach has been mainly studied by Viktor Lamme, Semir Zeki, David Milner and Melvyn Goodale. A second strategy, based on a more global neurophysiological approach, has been mainly developed by Ned Block, by Francis Crick and Christof Koch, and by Rodolfo Llinas, Antonio Damasio, Jean-Pierre Changeux and Gerald Edelman. In spite of some "overwhelming commonalities" among these theories, as Chalmers said it [17], nearly all of them have given a major role to interactions between the thalamus and the cortex.

### 4.2.1. First Strategy Focusing on the Visual System

**Model of Lamme.** The Local Recurrence theory of Viktor Lamme distinguished between three hierarchical types of neural processes related to consciousness [86-88]. The first stage involves a "feedforward sweep" during which the information is fed forward from striate visual regions (*i.e.*, V1) toward extrastriate areas as well as parietal and temporal cortices, without being accompanied by any conscious experience of the visual input. The second stage involves the "localized recurrent processing" during which the information is fed back to the early visual cortex. It is this recurrent interaction between early and

higher visual areas which leads to visual experience. The third and final stages consist of "widespread recurrent processing" which involves global interactions (as in the global workspace of Changeux) and extends toward the executive (*i.e.*, the frontoparietal network) and language areas. This final step also involves the attentional, executive, and linguistic processes that are necessary for conscious access and reportability of the stimulus. An interesting aspect of this theory is that it offers an explanation for the difference between conscious and unconscious perception in mechanistic rather in architectural terms. Another interesting aspect of this theory, although provocative, is that consciousness should not be defined by behavioral indexes such as the introspective reports of the subject. Instead, one should rely on neural indexes of consciousness, one of which is neural recurrence. Indeed, Lamme is concerned with the question of defining phenomenological consciousness when a report is impossible. However, one main difficulty with this theory is that it fails to take into account the recurrent connections that exist between regions that are not associated with consciousness (for instance between the area V1 and the thalamus). Although it remains possible that consciousness involves local recurrence between some cerebral components, this processus cannot then be considered as a sufficient condition for consciousness since it requires the involvement of additional mechanisms for explaining why it only applies to a restricted set of brain regions.

**Model of Zeki.** In the microconsciousness theory of Semir Zeki, it is assumed that instead of a single consciousness, there are multiple consciousness that are distributed in time and space [89,90]. This theory reflects the large functional specialization of the visual cortex. For example, while the perception of colors is associated with neural activity in area V4, motion perception reflects neural activity in MT/V5. Zeki took these findings as evidence that consciousness is not a unitary and singular phenomenon, but rather that it involves multiple consciousness that are distributed across processing sites, which are independent from each other. Another evidence in favor of this hypothesis is that the conscious perception of different attributes is not synchronous and can respect a temporal hierarchy. For example, psychophysical measures have shown that color is perceived a few tens of milliseconds before motion, reflecting the fact that neural activity during perception reaches V4 before reaching MT/V5. One main difficulty with this theory is that any processing region in the brain should, at first glance, constitute a neural correlates of consciousness in the multiple-consciousness theory. As such, it remains unclear why conscious perception is not associated with activity in most brain regions, including the cerebellum and subcortical regions, especially those con-

veying visual signals (e.g., the lateral geniculate nucleus). Another difficulty for this hypothesis is that visual regions can lead to the binding of several attributes in the absence of consciousness. This has been shown, for instance, in a patient with a bilateral parietal damage, suggesting that the binding mechanisms that are supposed to lead to microconsciousness can in fact operate in the absence of consciousness.

**Model of Milner and Goodale.** The duplex vision theory proposed by David Milner and Melvyn Goodale postulates that visual perception involves two interconnected, but distinctive pathways in the visual cortex, namely, the dorsal and the ventral stream [91,92]. After being conveyed along retinal and subcortical (*i.e.*, geniculate) structures, visual information reaches V1 and then involves two streams. The ventral stream projects toward the inferior part of the temporal cortex and serves to construct a conscious perceptual representation of objects, whereas the dorsal stream projects toward the posterior parietal cortex and mediates the control of actions directed at those objects. The two streams also differ at the computational and functional levels. On the one side, the ventral stream conveys information about the enduring (*i.e.*, long-lasting) characteristics that will be used to identify the objects correctly, and subsequently to link them to a meaning and classify them in relation to other elements of the visual scene. On the other side, the dorsal stream can be regarded as a fast and online visuomotor system dealing with the moment-to-moment information available to the system, which will be used to perform actions in real time. Recent evidence with visual masking has revealed unconscious neural activity in ventral regions, including the fusiform face area. This type of evidence is problematic for the duplex vision hypothesis since it predicts that conscious perceptions should be proportional to neural activity in the ventral system. Such a possibility of unconscious ventral processing can be nevertheless accommodated by assuming a threshold mechanism, as proposed in the model of Zeki [90]. However, including this threshold leads the theory to lose its former appeal, since consciousness is “only partially correlated” with activity in the ventral system.

#### 4.2.2. Second Strategy Based on a Global Neurophysiological Approach

**Model of Block.** One of the most influential issue in recent years has been the potential distinction between phenomenal and access consciousness proposed by Ned Block [93-95]. According to Block: “phenomenal consciousness is experience; the phenomenally conscious aspect of a state is what it is like to be in that state. The mark of access-consciousness, by contrast, is availability for use in reasoning and rationally guiding speech and

action”. In short, phenomenal consciousness results from sensory experiences such as hearing, smelling, tasting, and having pains. Block grouped together as phenomenal consciousness the experiences such as sensations, feelings, perceptions, thoughts, wants and emotions, whereas he excluded anything having to do with cognition, intentionality, or with properties definable in a computer program. In contrast, access consciousness is available for use in reasoning and for direct conscious control of action and speech. Access consciousness must be “representational” because only representational content can figure in reasoning. Examples of access consciousness are thoughts, beliefs, and desires. A point of controversy for this attempt to divide consciousness into phenomenal and access consciousness is that some people view the mind as resulting from fundamentally computational processes. This view of mind implies that all of consciousness is definable in a computer program. In fact, Block felt that phenomenal consciousness and access consciousness normally interact, but it is possible to have access consciousness without phenomenal consciousness. In particular, Block believed that zombies are possible and a robot could exist that is “computationally identical to a person” while having no phenomenal consciousness. Similarly, Block felt that one can have an animal with phenomenal consciousness but no access consciousness. If the distinction of Block between phenomenal and access consciousness is correct, then it has important implications for attempts by neuroscientists to identify the neural correlates of consciousness and for attempts by computer scientists to produce artificial consciousness in man-made devices such as robots. In particular, Block suggested that non-computational mechanisms for producing the subjective experiences of phenomenal consciousness must be found in order to account for the richness of human consciousness, or for there to be a way to rationally endow man-made machines with a similarly rich scope of personal experiences of “what it is like to be in conscious states”. However, many advocates of the idea that there is a fundamentally computational basis of mind felt that the phenomenal aspects of consciousness do not lie outside of the bounds of what can be accomplished by computation. Indeed, some of the conflict over the importance of the distinction between phenomenal consciousness and access consciousness centers on just what is meant by terms such as “computation”, “program” and “algorithm”, because one cannot know if it is within the power of “computation”, “program” or “algorithm” to produce human-like consciousness.

**Model of Llinas.** Rodolfo Llinas proposed a model in which the notion of emergent collective activity plays a central role [96-98]. He suggested that the brain is essentially a closed system capable of self-generated activity

based on the intrinsic electrical properties of its component neurons and their connectivity. For Llinas, consciousness arises from the ongoing dialogue between the cortex and the thalamus. The hypothesis that the brain is a closed system followed from the observation that the thalamus input from the cortex is larger than that from the peripheral sensory system, suggesting that thalamo-cortical iterative recurrent activity is the basis for consciousness. A crucial feature of this proposal was the precise temporal relations established by neurons in the cortico-thalamic loop. This temporal mapping was viewed as a functional geometry, and involved oscillatory activity at different scales, ranging from individual neurons to the cortical mantle. Oscillations that traverse the cortex in a highly spatially structured manner were therefore considered as candidates for the production of a temporal conjunction of rhythmic activity over large ensemble of neurons. Such gamma oscillations were believed to be sustained by a thalamo-cortical resonant circuit involving pyramidal neurons in layer IV of the neocortex, relay-thalamic neurons, and reticular-nucleus neurons. In this context, functional states such as wakefulness or sleep and other sleep stages are prominent examples of the breadth of variation that self-generated brain activity will yield. Since no gross morphological changes occur in the brain between wakefulness and dreamless sleep, the difference must be functional. Llinas proposed therefore that conscious processes of changes observed in the cycle sleep/dream/awakening rest on pairs of coupled oscillators, each pair connecting thalamus and the cortical region. He suggested also that temporal binding is generated by the conjunction of a specific circuit involving specific sensory and motor nuclei projecting to layer IV and the feedback via the reticular nucleus, and a non-specific circuit involving non-specific intralaminar nuclei projecting to the most superficial layer of the cortex and collaterals to the reticular and non-specific thalamic nuclei. Thus, the specific system was supposed to supply the content that relates to the external world, and the non-specific system was supposed to give rise to the temporal conjunction or the context. Together, they were considered as generating a single cognitive experience.

**Model of Crick and Koch.** In their framework, Francis Crick and Christof Koch tried to find the neural correlates of consciousness and suggested the existence of dynamic coalitions of neurons, in the form of neural assemblies whose sustained activity embodies the contents of consciousness [99-101]. By cortex they meant the regions closely associated with it, such as the thalamus and the claustrum. Crick and Koch began their theory with the notion of an “unconscious homunculus”, which is a system consisting of frontal regions of the brain “looking at the back, mostly sensory region”. They proposed that

we are not conscious of our thoughts, but only of sensory representations of them in imagination. The brain presents multiple rapid, transient, stereotyped and unconscious processing modules that act as “zombie” modes. This is in contrast to the conscious mode, that deals more slowly with broader, less stereotyped thoughts and responses. In this model, the cortex acts by forming temporary coalitions of neurons which support each other activity. The coalitions compete among each other, and the winning coalition represents what we are conscious of. These coalitions can vary in how widespread they are over the brain, and in how vivid and sustained they are. Moreover, more than one coalition can win at a given time. Especially, there might be different coalitions in the back and in the front of the cortex, where the coalitions in the front represent feelings such as happiness. An important point in this model is that consciousness may arise from certain oscillations in the cerebral cortex, which become synchronized as neurons fire 40 times per second. Crick and Koch believed the phenomenon might explain how different attributes of a single perceived object (its color and shape, for example), which are processed in different parts of the brain, are merged into a coherent whole. In this hypothesis, two pieces of information become bound together precisely when they are represented by synchronized neural firing. However, the extent and importance of this synchronized firing in the cortex is controversial. In particular, it remains a mystery: why should synchronized oscillations give rise to a visual experience, no matter how much integration is taking place? It should be added that, in this model, the neurons that are a part of the neural correlate of consciousness will influence many other neurons outside this correlate. These are called the penumbra, which exists of synaptic changes and firing rates. It also includes past associations, expected consequences of movements, and so on. It is by definition not conscious, but it might become so. Also, the penumbra might be the site of unconscious priming.

**Model of Damasio.** Antonio Damasio explored mainly the complexity of the human brain taking into consideration emotion, feeling, language and memory. According to him, the more important concepts are emotion, feeling, and feeling a feeling for the core consciousness [102,103]. The substrate for the representation of an emotional state is a collection of neural dispositions in the brain which are activated as a reaction on a certain stimulus. Once this occurs, it entails modification of both the body state and the brain state. This process starts as soon as the organism perceives, in the form of simple or complex sensory messages, proprioceptive or behavioral indicators meaning the danger or, on the contrary, the welfare. For Damasio, the emergence of feeling is based on the central role played by the proto-self that provides a map of

the state of the body [102]. The neuronal patterns which constitute the substrate of feeling arise in two classes of changes: changes related to body state, and changes related to cognitive state. Thus, a feeling emerges when the collection of neural patterns contributing to the emotion leads to mental images. The feelings correspond to the perception of a certain state of the body to which the perception of the corresponding state of mind is added, *i.e.* thoughts that the brain generates taking into account what it perceives of the state of the body. The notion of feeling is based on the organism detecting that the representation of its own body state has been changed by the occurrence of a certain object [102]. Consciousness is thus built on the basis of emotions transformed into feelings and feeling a feeling ; it mobilizes and gathers constantly, in a workspace, a certain number of informations necessary to strategies of survival and decision making. In this theory, consciousness is defined explicitly as a state of mind in which there is knowledge of one's own existence and of the existence of surroundings. According to Damasio, the emotions generate three types of levels of consciousness during the darwinian evolution: protoself, core self, and oneseif-autobiographical [102,104]. The protoself is the oneseif most primitive and most widespread within the alive species. It is constituted by special kinds of mental images of the body produced in body-mapping structures, below the level of the cerebral cortex. It results from the coherent but temporary interconnections of various cerebral charts of reentries of signals that represent the state of the organism at a given time. The protoself is an integrated collection of separate neural patterns that map moment by moment, the most stable aspects of the organism's physical structure. The first product of the protoself is primordial feelings. Whenever one is awake there has to be some form of feeling. On a higher level, the second form of the self, the core self, is about action. Damasio said that the core self unfolds in a sequence of images that described an object engaging the protoself and modifying that protoself, including its primordial feelings. These images are now conscious because they have encountered the self. On a still higher level there is the autobiographical self, constituted in large part by memories of facts and events about the self and about its social setting. It develops during social interactions to form what Damasio called "the wide consciousness". That is, the protoself and the core self constitute a "material me" whereas the autobiographical self constitutes a "social me". Our sense of person and identity is therefore in the autobiographical self.

All emotions engage structures related to the representation and/or regulation of organism state as, for example, the insular cortex, the secondary somatosensory cortex, the cingulated cortex, and nuclei in brainstem

tegmentum and hypothalamus [105]. These regions share a major feature in that they are all direct and indirect recipients of signals from the internal milieu, viscera and musculoskeletal frame. Damasio considered that no man can make decisions that are independent of the state of his body and his emotions. The influences that he undergoes are so numerous, that the assumption of a linear determinism which would determine him is not defensible. Till now, most theories had addressed emotion as a consequence of a decision rather than as the reaction arising directly from the decision itself. On the contrary, Damasio proposed that individuals make judgements not only by assessing the severity of outcomes and their probability of occurrence, but also primarily in terms of their emotional quality [106]. The key idea of Damasio was that decision making is a process that is influenced by signals that arise in bioregulatory processes, including those that express themselves in emotions and feelings. Decision making is not mediated by the orbito-frontal cortex alone, but arises from large-scale systems that include cortical and subcortical components. Like Dennett and Freeman, therefore, Damasio asserted that the individuals are under the influence of a chaotic causality of processes imprédictibles, indescribable, nonreproducible, but not random. Two main criticisms have been made about this theory [107]. First, Damasio tried to give an account of the mind as a set of unconscious mapping activities of the brain, and this did not presuppose, or at least it did not obviously presuppose, that these activities are conscious. But it is hard to understand any of these divisions of the self, protoself, core self and autobiographical self without supposing that they are already conscious. Second, Damasio stumbled over dreaming. Although phenomenal consciousness can be very vivid in dreams even when the rational processes of self-consciousness are much diminished, Damasio described dreams as mind processes unassisted by consciousness. He claimed that wakefulness is a necessary condition for consciousness. He described dreaming as paradoxical since, according to him, the mental processes in dreaming are not guided by a regular, properly functioning self of the kind we deploy when we reflect and deliberate. However, contrary to this point of view, it should be noted that dreaming is paradoxical only if one adopts a model of phenomenal consciousness based on self-consciousness (on knowledge, rationality, reflection and wakefulness).

**Model of Changeux.** The model of Changeux, developed with Stanislas Dehaene and collaborators [108,109], is founded on the idea that cerebral plasticity is mainly assured by a vast unit of interconnected neurons based on the model of "global workspace", historically related to Baar's cognitive model proposed in the 1980s [110]. Recall that the global workspace model of Baar is a process

which implies a pooling of information in a complex system of neural circuits, in order to solve problems that none of circuits could solve alone. This theory makes possible to explain how the consciousness is able to mobilize many unconscious processes, autonomous and localised in various cerebral regions (sensory, perceptuals, or imply in attention, evaluation, etc...), in order to treat them in a flexible and adjusted way [111]. Based on this workspace theory, Changeux proposed that neurons distributed with long-distance connections are capable of interconnecting multiple specialized processors. He also proposed that, when presented with a stimulus at threshold, workspace neurons can broadcast signals at the brain scale in an all-or-none manner, either highly activated or totally inactivated [108,112-114]. Workspace neurons thus allow many different processors to exchange information in a global and flexible manner. These neurons are assumed to be the targets of two different types of neuromodulatory inputs. First, they display a constantly fluctuating spontaneous activity, whose intensity is modulated by activating systems from cholinergic, noradrenergic and serotonergic nuclei in the brain stem, basal forebrain and hypothalamus. These systems modify the state of consciousness through different levels of arousal. Second, their activity is modulated by inputs arising from the limbic system, via connections to the anterior cingulate, orbitofrontal cortex, and the direct influence of dopaminergic inputs [115].

In this global workspace model, highly variable networks of neural cells are selected and their activity is reinforced by environmental stimuli. These reinforced networks can be said to 'represent' the stimuli, though no particular network of neural cells exclusively represents any one set of stimuli. The environment does not imprint precise images in memory. Rather, working through senses, the environment selects certain networks and reinforces the connections between them. These processes connect memory to the acquisition of knowledge and the testing of its validity. Every evocation of memory is a reconstruction on the basis of physical relative stable traces stored in the brain in latent form [116]. Changeux and Dehaene distinguished three kinds of neuronal representations: percepts; images, concepts, and intentions; and pre-representations. Percepts consist of correlated activity of neurons that is determined by the outer world and disintegrate as soon as external stimulation is terminated. Images, concepts, and intentions are activated objects of memory, which result from activating a stable memory trace (similar to the "remembered present" of Edelman). Pre-representations are multiple, spontaneously arising unstable; they are transient activity patterns that can be selected or eliminated [117]. In this line of idea, thoughts can be defined in terms of "calculations on the mental

objects" [118]. To be capable to be mobilized in the conscious workspace, a mental object must meet three criteria: 1) The object must be represented as a firing pattern of neurons; 2) The active workspace neurons must possess a sufficient number of reciprocal anatomical connections, particular in prefrontal, parietal and cingulate cortices; 3) At any moment, workspace neurons can only sustain a single global representation, the rest of workspace neurons being inhibited. This implies that only one active cortical representation will receive the appropriate top-down amplification and be mobilized into consciousness, the other representations being temporarily unconscious [119]. When a new situation happens, selection would occur from an abundance of spontaneously occurring pre-representations, namely those that are adequate to new circumstances and fit existing percepts and concepts. The free will is based on this adaptation process called 'resonance', because planning and decision making result from this selective adaptation process [117]. As in evolution, this could be a random recombination of representations. The meaning of representations involved would be given by their proper functions. Plans generated in this way would be intelligible because they could be appropriate for the situation. Naturally, linguistically coded representations could therein also be of central importance. It results from this global neurophysiological model that, as Changeux said it, the identity between mental states and physiological or physicochemical states of the brain is essential in all legitimacy and that it is enough with the man to be a neuronal man [118,120]. One important aspect that follows from this theory is that once a set of processors have stated to communicate through workspace connections, this multidomain processing stream also becomes more and more automatized through practice, resulting in direct interconnections, without requiring the use of workspace neurons and without involving consciousness. Another important aspect of this theory is that information computed in neural processors that do not possess direct long-range connections with workspace neurons, will always remain unconscious. This global workspace model has however been considered by some authors as a restrictive theory of access consciousness sacrificing some phenomenological aspects of consciousness. In other terms, it has been criticized for confounding the subjective experience of consciousness with the subsequent executive processes that are used to access its content.

**Model of Edelman.** To explain the primary consciousness, Edelman argued that living organisms must organize into perceptual categorization the unlabeled signals they receive from the environment. He started therefore his model by proposing a hypothesis which describes the fundamental process of categorization. This highly articulated hypothesis is mainly based on the theory of neu-

ronal group selection implying a basic process termed ‘reentry’, a massively recursive activity between neuronal groups. It is also based on the fundamental concept called “degeneracy”, whereby elements that are structurally different can perform the same function. The first idea of Edelman was to use the assumption of tying, imagined by Francis Crick for visual perception [22], to explain how the categories go from the form, color, texture and movement of external stimulations to the constitution of objectivity. His theory is based on the concept of neuronal chart. A chart is a sheet of neurons whose points are systematically connected, on the one hand, at points on sheets of receiving cells (skin, retina of the eye, etc...) to satisfy entries of signals and, on the other hand, at points located on other charts to satisfy reentries of signals functioning in the two directions. His second idea was to admit that the brain, genetically equipped at the birth with a superabundance of neurons, develops thanks to the death of a number of these neurons due to a process of darwinian selection. Lastly, his third idea was that parallel reentries of signals occur between the charts to ensure an adaptive answer. It is the ongoing recursive dynamic interchange of signals occurring in parallel between maps that continually interrelates these maps to each other in space and time. Categorization, therefore, is ensured by the dynamic network of a considerable number of entries and reentries selected by a multitude of charts [26,121,122]. Although distributed on many surfaces of the brain, the inputs, at the same time connected between them and connected to charts of reentries, can thus react the ones on the others in a mechanism of functional segregation for finally giving a unified representation of the object. In other words, categorization is a global “encartage” which does not need a program *a priori*, a “homoncule”, to produce it. For Edelman, this mechanism is a plausible working hypothesis because his research group has shown that brain-based devices are capable of perceptual categorization and conditioned behavior [123]. Over the last 14 years, a series of robotic organisms (called Darwin series) have indeed been used to study perception, operant conditioning, episodic and spatial memory, and motor control through the simulation of brain regions such as the visual cortex, the dopaminergic reward system, the hippocampus, and the cerebellum [124]. However, to reach the major stage of conscious perceptive experience, Edelman thought that at least six further conditions should be met [26,122,125,126]. Firstly, the brain must have memory. Memory, according to Edelman, is an active process of recategorization on the basis of a former categorization. The perceptible category, for example, is acquired by an experience by means of charts of reentries. But if there is a perceptive entry by seeing, for example, a table again, one recategorizes this entry by improving preceding ca-

tegorization, and so on; one reinvents the category continuously. In other terms, memory generates “information” by construction: it is creative but nonreplicative. Secondly, the brain must have a system of training. The training implies changes of behavior which rest on categorization controls by positive or negative values; an animal can choose, for example, what is clear or dark, hot or cold, etc, to satisfy its needs for survival. Thirdly, although self-awareness appears only with the higher consciousness, the brain must have a system of diffuse feelings which enables it to discriminate between oneself and not oneself. Fourthly, the brain must be able to categorize the succession of events in time and to form concepts pre-linguistics. Edelman thought that the categorization of successive events and the formation of these concepts have a common neurobiological substrate. Fifthly, there must be continuous interactions between conditions 1, 2, 3 and 4 so that a particular memorial system for the values is formed. Sixthly, the brain needs also to have reentrant connections between this particular memorial system and the anatomical systems. That is, according to Edelman, to formulate a model of primary consciousness it is necessary to have the notions of perceptual categorization, concept formation, and value-category memory in hand.

As Changeux, Edelman considered that it is the thalamo-cortical system, called “dynamic core”, which contributes directly to the conscious experience [125,126]. It is a dense fibre network inter-connected in the two directions, in which connections occur unceasingly and come in competition so that the most effective fibres carry it (according to a process of darwinian competition). This reentrant dynamic core is thus a process, not a substance. It is in perpetual rebuilding (with periods of time of a few hundred milliseconds), in wide space but formed nevertheless of unified and highly integrated elements. As spatial process, it ensures the coherence of consciousness states; it is the principal dynamic driving core of the constitution of self. The key structures for the emergence of consciousness are the specific thalamic nuclei, modulated by the reticular nucleus in their reentrant connectivity to cortex, the intralaminar nuclei of the thalamus, and the grand system of corticocortical fibers. What emerges from these interactions is the ability to construct a scene. This activity is influenced by animal’s history of rewards and punishments accumulated during its past. It is essential for constructing, within a time window of fractions of seconds up to several seconds, the particular memory of animals, that is the “remembered present” [126]. Higherorder consciousness is consciousness to be conscious, self-awareness. It is accompanied by a sense of self and the ability to construct past and future scenes in the waking state. It is the activity of the dynamic core which converts the signals received from the outside and

inside of the body into oneself, what Edelman called “phenomenal transformation” [125]. This higher consciousness requires, at the minimum, a semantic capability and, in its most developed form, a linguistic capability which makes possible to symbolize the relations of the past, present and future and to do representations. On the other hand, Edelman asserted that the linkage proceeds from the underlying neural activity to the conscious process and not the reverse, but that consciousness is not directly caused by the neural processes, it accompanies them. Rather, they are the neural processes and the entire body which are causal. Consciousness is “reincorporated” in the body, insofar as it is one of the modulations by which the body expresses its decisions; it amplifies the effects of the body to produce the free will [125]. Since the body “is decided” by many nonlinear and chaotic determinisms (described by Alain Berthoz [127]), consciousness is thus also “decided” by these complex body determinations. However, consciousness adds to the body an additional level of complexity, because it is able to make differences within a vast repertory of informations. What makes indeed informative a state of human consciousness is the fact that the man is able to make a difference between billion states different of things. The intentional understanding of consciousness to reduce uncertainty is always singular and gives the character subjective and private to mental-lived of each individual at every moment. Like Dennett, Freeman and Damasio, Edelman thus admitted that the free will is a nonlinear chaotic causality, generally circular, which results from the complex interactions between the body and the brain in the local environmental contexts. The model of Edelman offers an interesting tentative of unifying the hard and easy problems. Indeed, since the reentrant dynamic core theory considers consciousness as an emergent property of any system sharing specific core mechanisms of differentiated integration, the most reliable index of consciousness should reflect the quantification of this core mechanism. However, a conceptual difficulty of this hypothesis lies in the paradox of trying to prove that neural indexes are more respectable because they supposedly probe phenomenal and not access consciousness. Indeed, neural indexes of any sort have to be validated by confronting them at one point with some kind of report, hence with access and not phenomenal consciousness. Even if it turns out that one of these neural indexes turns out to be perfectly correlated with consciousness, and thus becomes a perfectly reliable measure of consciousness, then one might still ask whether we have made any progress.

## 5. Conclusions

A great variety of specific theories have been proposed,

whether cognitive, phenomenal, physics, neural or based on Artificial Intelligence, to explain consciousness as a natural feature of the physical world. The most prominent examples of philosophical and non-neural approaches included the naturalistic dualism model of David Chalmers, the multiple draft cognitive model of Daniel Dennett, the phenomenological theory of Francisco Varela and Humberto Maturana and the physics-based approaches of Stuart Hameroff and Roger Penrose. The major neurophysiological theories included the emerging spatiotemporal patterns of Walter Freeman, the visual system-based approaches of Viktor Lamme, Semir Zeki, David Milner and Melvyn Goodale, the phenomenal/access consciousness model of Ned Block, the synchronous cortico-thalamic rhythms of Rodolfo Llinas, the synchronized firing cortical oscillations of Francis Crick and Christof Koch, the emotive somatosensory processes of Antonio Damasio, the neuronal global workspace of Jean-Pierre Changeux and Stanislas Dehaene, and the reentrant cortical loops of Gerald Edelman and Giulio Tononi. It is foreseeable that the range of these models will extend in the future. For example, it is supposed now that consciousness associated with the dream would be a rebuilding at the time of the waking, related to relaxations of specific neuromodulators on very short intervals of time [128]. Such a hypothesis might generate new models. On the other hand, new perspectives of Artificial Intelligence have attempt to check the hypothesis of the functioning of brain architecture and simulate consciousness using both the methods and laws of informational machines and the processing capabilities of biological systems. For example, series of brain-based robots have been tested over the last decade to study perception, operant conditioning, episodic and spatial memory, and motor control. Moreover, a supercomputer program, called Human Brain Project, was given for objective to attempt to reproduce an artificial consciousness in the future.

One can be skeptic, like Paul Ricoeur, as for the possibility of obtaining a true theory which makes a complete synthesis between a speech “neuro” and a “psychological” speech. The body is this object which is both *me* and *mien*, but there is no passage from one speech to another [120]. Any attempt to confuse the figures of the body-object and body-lived of consciousness, was it founded on the principles and methods of neuroscientific approaches more innovating, is obviously impossible for ontological reasons.

However, like Edelman says it, to deplore that one cannot build a scientific model of consciousness without being able to theoretically explain the quality of a quale (what one feels, for example, by seeing the red of a flower) is a non-problem; it is a problem of the same order as to want to answer the famous philosophical question: “why there is something and why not rather nothing?” [46].

The subjective aspect of consciousness is a pure obviousness of the real world. It is a pure sensitive which is given of only one blow of the medium of itself [129], which it is impossible to analyze it without causing its disappearance immediately. It does not remain about it less than the neurobiological understanding of relationships between the brain and the objective aspects of consciousness is the great challenge which will animate researches of future decades. Certainly, the phenomenological model of Varela is compatible with some experimental neurobiological approaches. Likewise, the physics-based model of Hameroff and Penrose could be open to some researches of mental-brain correlations via the study of signaling mechanisms of synaptic functions. Moreover, Artificial Intelligence has proven to be a fertile physical enterprise to check some hypothesis about the functioning of brain architecture, although no machine has been capable of reproducing an artificial consciousness. However, only the advances in neurosciences seems actually capable of taking into account the constraints of a wide range of complex properties to provide an adequate explanatory basis for consciousness. Of these properties, several stand out as particular challenges to theoretical effort, in particular the property of intentionality, the problem of explaining both the first-person and third-person aspects of consciousness, and the question of plausible necessary preconditions to develop self-awareness. As attests it analysis and the myriad of neurophysiological data already available, the states (arousal) of consciousness, the behavior, the brain processes of conscious systems, and a number of subjective contents of consciousness (some qualia) are objectivable and can be experimentally studied. In spite of innumerable difficulties which one will still meet to study and understand a phenomenon which is the originating base of categorical representations of the intersubjective and subjective world, it is therefore possible that one will manage one day to build a naturalistic satisfying explanation of consciousness that links together molecular, neuronal, behavioral, psychological and subjective data in an unified, coherent scientific theory.

## 6. References

- [1] R. Descartes, "The Philosophical Works of Descartes," Cambridge University Press, Cambridge, 1975.
- [2] P. H. Nidditch, "Enquiries Concerning the Human Understanding," 3rd Edition, Clarendon Press, Oxford, 1975.
- [3] J.-N. Missa, "Que Peut-On Espérer d'Une Théorie Neuroscientifique de la Conscience? Plaidoyer Pour une Approche Evolutionniste," In: P. Poirier and L. Faucher, Eds., *Des neurosciences à la philosophie. Neurophilosophie et philosophie des neurosciences*, Syllepse, Paris, 2008, pp. 356-357.
- [4] S. Nicolas, "Psychologie de W. Wundt," L'Harmattan, Paris, 2003.
- [5] R. Rudio, "William James, Philosophie, Psychologie Religion," L'Harmattan, Paris, 2008.
- [6] G. Schiemann, "Hermann von Helmholtz's Mechanism: The Loss of Certainty. A Study on the Transition from Classical to Modern Philosophy of Nature," Springer, Dordrecht, 2009.
- [7] E. Husserl, "Idées Directrices Pour une Phénoménologie," Gallimard, Paris, 1950.
- [8] J. Watson, "Behaviorism," W. W. Norton, New York, 1924.
- [9] B. F. Skinner, "Science and Human Behaviour," Mac Millan, New York, 1953.
- [10] U. Neisser, "Cognitive Psychology," Prentice Hall, Englewood Cliffs, 1965.
- [11] H. Gardner, "Frames of Mind: The Theory of Multiple Intelligences," Basic Books, New York, 1983.
- [12] M. Heidelberger, "Archived Draft of the Mind-Body Problem in the Origin of Logical Empirism: Herbert Feigl and Psychophysical Parallelism," Pittsburg University Press, Pittsburg, 2003.
- [13] D. Robinson, "Epiphenomenalism, Laws and Properties," *Philosophical Studies*, Vol. 69, No. 1, 1982, pp. 1-34. [doi:10.1007/BF00989622](https://doi.org/10.1007/BF00989622)
- [14] P. S. Churchland, "Que Peut Nous Enseigner la Neurobiologie au Sujet de la Conscience?" Poirier Faucher, 2008, pp. 329-354.
- [15] J. M. Roy, "Naturalism Emergentist and Causal Explanation," *Intellectica*, Vol. 39, 2004, pp. 199-227.
- [16] D. J. Chalmers, "Facing Up to the Problem of Consciousness," *Journal of Consciousness Studies*, Vol. 2, 1995, pp. 200-219.
- [17] D. J. Chalmers, "The Conscious Mind: In Search of a Fundamental Theory," Oxford University Press, Oxford, 1996.
- [18] C. Mc Ginn, "Consciousness and Its Objects," Oxford University Press, Oxford, 2004.
- [19] T. Nagel, "What Is It Like to Be a Bat?" *Philosophical Review*, Vol. 83, No. 4, 1974, pp. 435-450. [doi:10.2307/2183914](https://doi.org/10.2307/2183914)
- [20] D.C. Dennett, "Quining Qualia," In: A. Marcel and E. Bisiach, Eds., *Consciousness in Contemporary Science*, Oxford University Press, Oxford, 1988, pp 42-77.
- [21] D. C. Dennett and M. Kinsbourne, "Time and the Observer: The Where and When of Consciousness in the Brain," *Behavioral and Brain Sciences*, Vol. 15, No. 2, 1992, pp. 183-247. [doi:10.1017/S0140525X00068229](https://doi.org/10.1017/S0140525X00068229)
- [22] J. R. Searle, "Le Mystère de la Conscience," Odile Jacob, Paris, 1999.
- [23] D. C. Dennett, "Freedom Evolves," Viking Press, New York, 2003.
- [24] D. C. Dennett, "Darwin's Dangerous Idea: Evolution and the Meanings of Life," Simon & Schuster, New York, 1996.



- [25] R. Dawkin, "The Selfish Gene," 2nd Edition, Oxford University Press, Oxford, 1989.
- [26] G. M. Edelman, "Bright Air, Brilliant Fire: On the Matter of Mind," The Penguin Press, London, 1992.
- [27] D. J. Chalmers, "The Cognitive Neurosciences III," MIT Press, Cambridge, 2004.
- [28] C. I. Lewis, "Mind and the World Order," Charles Scribners Sons, New York, 1929.
- [29] B. Keeley, "The Early History of the Quale and Its Relation to the Senses," In: J. Symons and P. Calvo, Eds., *Routledge Companion to the Philosophy of Psychology*, Routledge Press, New York, 2009.
- [30] J. Levine, "Qualia: Intrinsic, Relational or What?" In: T. Metzinger, Ed., *Conscious Experience*, Schöningh, Paderborn, 1995, pp. 277-292.
- [31] A. N. Whitehead, "Process and Reality: An Essay in Cosmology," Macmillan, New York, 1929.
- [32] J. Haugeland, "Artificial Intelligence: The Very Idea," The MIT Press, Cambridge, 1985.
- [33] C. Koch, "The Quest for Consciousness: Neurobiological Approach," Roberts, Englewood, 2004.
- [34] T. Horgan and J. Tienson, "The Intentionality of Phenomenology and Phenomenology of Intentionality," In: D. Chalmers, Ed., *Philosophy Ofmind: Classical and Contemporary Readings*, Oxford University Press, 2002, pp. 520-533.
- [35] J. Delacour, "Conscience and Brain. The New Border of the Neurosciences," DeBoeck University, Brussels, 2001.
- [36] J. Delacour, "Biology of the Conscience," *Philosophical Review of France and Abroad*, Vol. 129, 2004, pp. 315-332.
- [37] J. Hohwy, "The Neural Correlates of Consciousness: New Experimental Approaches Needed?" *Consciousness and Cognition*, Vol. 18, No. 2, 2009, pp. 428-438. [doi:10.1016/j.concog.2009.02.006](https://doi.org/10.1016/j.concog.2009.02.006)
- [38] J. Delacour, "Neurobiology of Consciousness: An Overview," *Behavioural Brain Research*, Vol. 85, No. 2, 1997, pp. 127-141. [doi:10.1016/S0166-4328\(96\)00161-1](https://doi.org/10.1016/S0166-4328(96)00161-1)
- [39] T. J. Balkin, A. L. Braun, N. J. Wassenstein, K. Jeffries, M. Varga, P. Baldwin, G. Belenky and P. Herscovitch, "The Process of Awakening: A PET Study of Regional Brain Activity Patterns Mediating the Re-Establishment of Alertness and Consciousness," *Brain*, Vol. 125, No.10, 2002, pp. 2308-2319. [doi:10.1093/brain/awf228](https://doi.org/10.1093/brain/awf228)
- [40] D. J. Chalmers, "What Is a Neural Correlate of consciousness?" In: T. Metzinger, Ed., *Neural Correlates of Consciousness: Empirical and Conceptual Questions*, MIT Press, Cambridge, 2000.
- [41] G. M. Edelman, "Naturalizing Consciousness: A Theoretical Framework," *Proceedings of the National Academy of Sciences USA*, Vol. 100, No. 9, 2003, pp. 5520-5524. [doi:10.1073/pnas.0931349100](https://doi.org/10.1073/pnas.0931349100)
- [42] G. Tononi and C. Koch, "The Neural Correlates Ofconsciousness: An Update," *Annals of the New York Academy of Sciences*, Vol. 1124, No. 1, 2008, pp. 261-298. [doi:10.1196/annals.1440.004](https://doi.org/10.1196/annals.1440.004)
- [43] C.-Y. Kim and R. Blake, "Psychophysical Magic: Rendering the Visible 'Invisible'," *Trends in Cognitive Sciences*, Vol. 9, No. 8, 2004, 381-388. [doi:10.1016/j.tics.2005.06.012](https://doi.org/10.1016/j.tics.2005.06.012)
- [44] D. G. Edelman and A. K. Seth, "Animal Consciousness: A Synthetic Approach," *Trends in Neurosciences*, Vol. 32, No. 9, 2009, pp. 476-484. [doi:10.1016/j.tins.2009.05.008](https://doi.org/10.1016/j.tins.2009.05.008)
- [45] J. A. Mather, "Cephalopod Consciousness: Behavioural Evidence," *Consciousness and Cognition*, Vol. 17, No. 1, 2008, pp. 37-48. [doi:10.1016/j.concog.2006.11.006](https://doi.org/10.1016/j.concog.2006.11.006)
- [46] B. Lechevalier, F. Eustace and F. Viader, "Conscience and Its Disorders," De Boeck University, Brussels, 1998, pp. 151-166.
- [47] G. Buttazzo, "Can a Machine Ever Become Self-Aware? Artificial Humans," Goethe Institute, Los Angeles, 2000, pp. 45-49.
- [48] G. G. Towell and J. W. Shavlik, "Knowledge-Base Artificial Neural Networks," *Artificial Intelligence*, Vol. 70, No. 1-2, 1994, pp. 119-165. [doi:10.1016/0004-3702\(94\)90105-8](https://doi.org/10.1016/0004-3702(94)90105-8)
- [49] R. Andrews, J. Diederich and A. B. Tickle, "Survey an Critique of Techniques for Extracting Rules from Trained Artificial Neural Networks," *Knowledge-Based System*, Vol. 8, No. 6, 1995, pp. 373-389. [doi:10.1016/0950-7051\(96\)81920-4](https://doi.org/10.1016/0950-7051(96)81920-4)
- [50] P. Stern and J. Travis, "Of Bytes and Brains," *Science*, Vol. 314, No. 5796, 2006, pp. 75-77. [doi:10.1126/science.314.5796.75](https://doi.org/10.1126/science.314.5796.75)
- [51] D. E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning," Addison-Wesley, Reading, 1989.
- [52] Z. Michalewicz, "Genetic Algorithm, Data Structures, Evolution Programs," 3rd Edition, Springer-Verlag, New York, 1996.
- [53] A. E. Eiben and M. Schoenauer, "Evolutionary Computing," *Information Processing Letters*, Vol. 82, No. 1, 2002, pp. 1-6. [doi:10.1016/S0020-0190\(02\)00204-1](https://doi.org/10.1016/S0020-0190(02)00204-1)
- [54] E. Bonabeau, "Agent-Based Modeling: Methods and Techniques for Simulating Human Systems," *Proceedings of the National Academy of Sciences USA*, Vol. 99, No. 3, 2002, pp. 7280-7287. [doi:10.1073/pnas.082080899](https://doi.org/10.1073/pnas.082080899)
- [55] A. Cardon, "A Multi-Agent Model for Co-Operative Communications in Crisis Management Systems: The Act of Communication," *Information Modeling and Knowledge Bases*, IOS Press, Amsterdam, 1998, pp. 66-82.
- [56] A. Cardon, T. Galinho and J.-P. Vacher, "Genetic Algorithm Using Multi-Objective in a Multi-Agent System," *Robotic and Autonomous Systems*, Vol. 33, 2000, pp. 179-190.
- [57] A. Cardon, "Design and Behaviour of a Massive Organization of Agents," *Design of Intelligent Multi-Agent Systems, Human-Centredness, Architectures, Learning and Adaptation*, Springer, Berlin/Heidelberg, Vol. 162, 2004, pp. 133-190.

- [58] A. Cardon, "Artificial Consciousness, Artificial Emotions, and Autonomous Robots," Springer, Berlin/Heidelberg, 2006.
- [59] R. Sanz, I. Lopez and J. Bermejo-Alonso, "A Rational and Vision for Machine Consciousness in Complex Controllers" *Artificial Consciousness*, Imprint Academic, Exeter, 2007.
- [60] Y. Shoham and K. Leyton-Brown, "Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations," Cambridge University Press, Cambridge, 2009.
- [61] J. L. Krichmar and G.M. Edelman, "Machine Psychology: Autonomous Behavior, Perceptual Categorization, and Conditioning in a Brain-Based Device" *Cerebral Cortex*, Vol. 12, No. 8, 2002, pp. 818-830.  
[doi:10.1093/cercor/12.8.818](https://doi.org/10.1093/cercor/12.8.818)
- [62] O. Holland, R. Knight and R. Newcombe, "The Role of the Self Process in Embodied Machine Consciousness," *Artificial Consciousness*, Imprint Academic, Exeter, 2007.
- [63] H. Markram, "The Blue Brain Project," *Nature Reviews Neuroscience*, Vol. 7, 2006, pp. 153-160.  
[doi:10.1038/nrn1848](https://doi.org/10.1038/nrn1848)
- [64] D. Hofstadter and D. Dennett, "Sights of the Spirit," InterEditions, Paris, 1987.
- [65] G. Hesslow and D. A. Jirenhed, "Must Machines Be Zombie? Internal Simulation as a Mechanism for Machine Consciousness," AAA Symposium, Washington DC, 8-11 November 2007, pp. 1-6.
- [66] H. Maturana and F. Varela, "Autopoiesis and Cognition: The Realization of the Living," D. Reidel Publishing Co., Dordrecht, 1980. [doi:10.1007/978-94-009-8947-4](https://doi.org/10.1007/978-94-009-8947-4)
- [67] F. J. Varela, "Neurophenomenology: A Methodological Remedy for the Hard Problem," *Journal of Consciousness Studies*, Vol. 3, 1996, pp. 330-349.
- [68] F. J. Varela, "The Naturalization of Phenomenology as the Transcendence of Nature: Searching for Mutual Generative Constraints," *Alter: Revue de Phénoménologie*, Vol. 5, 1997, pp. 355-385.
- [69] F. J. Varela, "A Science of Consciousness as If Experience Mattered," In: S. Hameroff, A. W. Kaszniak and A. C. Scott, Eds., *Towards a Science of Consciousness II: The Second Tucson Discussions and Debates*, MIT Press, Cambridge, 1998.
- [70] D. Rudrauf, A. Lutz, D. Cosmelli, J.-P. Lachaux and M. LeVan Quyen, "From Autopoiesis to Neurophenomenology: Francisco Varela's Exploration of the Biophysics of Being," *Biological Research*, Vol. 36, 2003 pp. 1-28.
- [71] M. Merleau-Ponty, "The Structure of Behavior," Beacon Press, Boston, 1963.
- [72] F. J. Varela, "Organism: Ameshworkof Selfless Selves," *Organism and the Origin of Self*, Kluwer, Dordrecht, 1991, pp. 79-107.
- [73] F. J. Varela, "Steps to a Science of Interbeing: Unfolding Implicit the Dharma in Modern Cognitive Science," In: S. Bachelard, G. Claxton and G. Watson, Eds., *The Psychology of Awakening: Buddhism, Science and Our Day to Day Lives*, Rider/Randol House, New York, 1999, pp. 71-89.
- [74] P. L. Luisi, "Autopoiesis: A Review and a Reappraisal," *Naturwissenschaften*, Vol. 90, 2003, pp. 49-59.
- [75] R. Penrose, "The Emperor's New Mind," Oxford University Press, Oxford, 1989.
- [76] R. Penrose, "Shadows of the Mind: A Search for the Missing Science of Consciousness," Oxford University Press, Oxford, 1994.
- [77] S. Hameroff and R. Penrose, "Orchestrated Reduction of Quantum Coherence in Brain Microtubules: A Model for Consciousness," *Neural Network World*, Vol. 5, 1995, pp. 793-804.
- [78] S. Hameroff and R. Penrose, "Orchestrated Reduction of Quantum Coherence in Brain Microtubules: A Model for Consciousness," In: S. R. Hameroff and A. C. Scott, Eds., *Toward a Science of Consciousness, The First Tucson Discussions and Debates*, MIT Press, Cambridge, 1996.
- [79] R. Grush and P. Churchland, "Gaps in Penrose's Toiling," *Journal of Consciousness Studies*, Vol. 2, 1995, pp. 10-29.
- [80] W. J. Freeman, "Mass Action in the Nervous System," Elsevier Science & Technology Books, New York, 1975.
- [81] W. J. Freeman, "How Brains Make Up to Their Minds," Weidenfeld, London, 1999.
- [82] W. J. Freeman, "Neurodynamics; Exploration of Mesoscopic Brain Dynamics," Springer-Verlag, London, 2000.  
[doi:10.1007/978-1-4471-0371-4](https://doi.org/10.1007/978-1-4471-0371-4)
- [83] W. J. Freeman, "Mass Action in the Nervous System," 2004. <http://sulcus.berkeley.edu/>
- [84] W. J. Freeman "Origin, Structure and Role of Background EEG Activity, Part 3, Neural Frame Simulation," *Clinical Neurophysiology*, Vol. 116, No. 5, 2006, pp. 1118-1129. [doi:10.1016/j.clinph.2004.12.023](https://doi.org/10.1016/j.clinph.2004.12.023)
- [85] J. J. Wright, "Cortical Phase Transitions: Properties Demonstrated in Continuum Simulations at Mesoscopic and Macroscopic Scales," *Journal of New Mathematics and Natural Computation*. Vol. 5, No. 1, 2009, pp. 159-193.
- [86] V. A. F. Lamme and P. R. Roelfsema, "The Distinct Modes of Vision Offered by feedforward and Recurrent Processing," *Trends in Neurosciences*, Vol. 23, No. 11, 2000, pp. 571-579.  
[doi:10.1016/S0166-2236\(00\)01657-X](https://doi.org/10.1016/S0166-2236(00)01657-X)
- [87] V. A. F. Lamme, "Why Visual Attention and Awareness Are Different," *Trends in Cognitive Sciences*, Vol. 7, No. 1, 2003, pp. 12-18. [doi:10.1016/S1364-6613\(02\)00013-X](https://doi.org/10.1016/S1364-6613(02)00013-X)
- [88] V. A. F. Lamme, "Towards a True Neural Stance on Consciousness," *Trends in Cognitive Sciences*, Vol. 10, No. 11, 2006, pp. 494-501. [doi:10.1016/j.tics.2006.09.001](https://doi.org/10.1016/j.tics.2006.09.001)
- [89] S. Zeki, "A Vision of the Brain," Blackwell Scientific Publications, Oxford, 1993.
- [90] S. Zeki, "Inner Vision: An Exploration of Art and the Brain," Oxford University Press, Oxford, 1999.
- [91] A. D. Milner and M. A. Goodale, "The Visual Brain in Action," Oxford University Press, Oxford, 1998.
- [92] M. A. Goodale and A. D. Milner, "Sight Unseen: An Exploration of Conscious and Unconscious Vision," Ox-

- ford University Press, Oxford, 2004, p. 135.
- [93] N. Block, "How Can We Find the Neural Correlate of Consciousness?" *Trends in Neurosciences*, Vol. 19, 1996, pp. 456-459.
- [94] N. Block, "On a Confusion About a Function of Consciousness," In: N. Block, O. Flanagan and G. Güzeldere, Eds., *The Nature of Consciousness: Philosophical Debates*, The MIT Press, Cambridge, 1997, pp. 375-416.
- [95] N. Block, "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience," *Behavioral and Brain Sciences*, Vol. 30, No. 5-6, 2007, pp. 481-548. doi:10.1017/S0140525X07002786
- [96] R. Llinas and D. Pare, "Of Dreaming and Wakefulness," *Neuroscience*, Vol. 44, No. 3, 1991, pp. 521-535. doi:10.1016/0306-4522(91)90075-Y
- [97] R. Llinas, U. Ribary, D. Contreras and C. Pedroarena, "The Neuronal Basis for Consciousness," *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 353, No. 1377, 1998, pp. 1841-1849. doi:10.1098/rstb.1998.0336
- [98] R. Llinas and U. Ribary, "Consciousness and the Brain. The Thalamocortical Dialogue in Health and Disease," *Annals of the New York Academy of Sciences*, Vol. 929, 2001, pp. 166-175.
- [99] F. C. Crick and C. Koch, "Towards a Neurobiological Theory of Consciousness," *Seminars in Neuroscience*, Vol. 2, 1990, pp. 263-275.
- [100] F. C. Crick and C. Koch, "Consciousness and Neuroscience," *Cerebral Cortex*, Vol. 8, No. 2, 1998, pp. 97-107. doi:10.1093/cercor/8.2.97
- [101] F. C. Crick and C. Koch, "A Framework for Consciousness," *Nature Reviews Neuroscience*, Vol. 6, 2003, pp. 119-126.
- [102] A. R. Damasio, "The Feeling of What Happens: Body and Emotion in the Making of Consciousness," Harcour Brace & Co, New York, 1999.
- [103] A. R. Damasio, "Looking for Spinoza: Joy, Sorrow, and the Feeling Brain," Harcourt Inc, Orland, 2003.
- [104] A. R. Damasio, "Descartes's Error," Avon Books, New York, 1994.
- [105] A. Damasio, T. Ji. Grabowsk, A. Bechara, H. Damasio, L. L. B. Ponto, J. Parvizi and R. D. Hichwa, "Subcortical and Cortical Brain Activity during the Feeling of Self-Generated Emotions," *Nature Reviews Neuroscience*, Vol. 3, 2000, pp. 1049-1056.
- [106] A. Bechara, H. Damasio and A. R. Damasio, "Emotion, Decision Making and the Orbitofrontal Cortex," *Cerebral Cortex*, Vol. 10, No. 3, 2000, pp. 295-307. doi:10.1093/cercor/10.3.295
- [107] J. R. Searle, "The Mystery of Consciousness Continues," The New York Review of Books, 2000, <http://www.nybooks.com/articles/archives/2011/jun/09/mystery-consciousness-continues>
- [108] S. Dehaene, M. Kerszberg and J.-P. Changeux, "A Neuronal Model of a Global Workspace in Effortful Cognitive Tasks," *Proceedings of the National Academy of Sciences USA*, Vol. 95, No. 24, 1998, pp. 14529-14534. doi:10.1073/pnas.95.24.14529
- [109] S. Dehaene, J.-P. Changeux, L. Naccache, J. Sackur and C. Sergeant, "Conscious, Preconscious, and Subliminal Processing: Has Testable Taxonomy," *Trends in Cognitive Sciences*, Vol. 10, 2006, pp. 204-211.
- [110] B. J. Baars, T. Z. Ramsay and S. Laureys, "Brain Conscious Experience and the Observing Self," *Trends in Cognitive Sciences*, Vol. 26, No. 12, 2003, pp. 671-675. doi:10.1016/j.tins.2003.09.015
- [111] B. J. Baar, "A Cognitive Theory of Consciousness," Cambridge University Press, Cambridge, 1988.
- [112] S. Dehaene and J.-P. Changeux, "Reward-Dependent Learning in Neuronal Networks for Planning and Decision Making," *Progress in Brain Research*, Vol. 126, 2000, pp. 217-229. doi:10.1016/S0079-6123(00)26016-0
- [113] S. Dehaene and L. Naccache, "Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework," *Cognition*, Vol. 79, No. 1-2, 2001, pp. 1-37. doi:10.1016/S0010-0277(00)00123-2
- [114] S. Dehaene, C. Sergent and J.-P. Changeux, "A Neuronal Network Model Linking Subjective Reports and Objective Physiological Data during Conscious Perception," *Proceedings of the National Academy of Sciences USA*, Vol. 100, No. 14, 2003, pp. 8520-8525. doi:10.1073/pnas.1332574100
- [115] J.-P. Changeux and S. Dehaene, "The Workspace Model: Conscious Processing and Learning," In: R. Menzel, Ed., *Learning Theory and Behavior, Vol 1 of Learning and Memory: A Comprehensive Reference*, Elsevier, Oxford, 2008, pp. 729-758.
- [116] J.-P. Changeux, "The Physiology of Truth," Harvard University Press, Cambridge, 2004.
- [117] H. Walter, "Neurophilosophy of Free Will. Frontal Cortex and Intelligibility," *The Oxford Handbook of Free Will*, 2002, pp. 565-570.
- [118] J.-P. Changeux, "The Neuronal Man: The Biology of Mind," Princeton University Press, Princeton, 1985.
- [119] S. Dehaene and J.-P. Changeux, "Neural Mechanisms for Access to Consciousness," *The Cognitive Neurosciences III*, MIT Press, Cambridge, 2004.
- [120] J.-P. Changeux and P. Ricoeur, "What Makes Us Think," Princeton University Press, Princeton, 2002.
- [121] G. Tononi, O. Sporns and G. M. Edelman, "Reentry and The Problem of Integrating Multiple Cortical Areas: Simulation of Dynamic Integration in the Visual System," *Cerebral Cortex*, Vol. 2, No. 4, 1992, pp. 310-335. doi:10.1093/cercor/2.4.310
- [122] G. M. Edelman and G. Tononi, "A Universe of Consciousness: How Matter Becomes Imagination," In: G. M. Edelman and G. Tononi, Eds., *Basic Books*, New York, 2000.
- [123] G. N. Reeke, L. H. Finkel, O. Sporns and G. M. Edelman, "Synthetic Neural Modeling: A Multilevel Approach to the Analysis of Brain Complexity," *Bulletin of the French*

- Company of Philosophy*, Conference of February 28, 1990, pp. 607-707.
- [124] J. L. Krichmar and G. M. Edelman, "Brain-Based Devices for the Study of Nervous Systems and the Development of Intelligent Machines," *Artificial Life*, Vol. 11, No. 1-2, 2005, pp. 63-77.  
[doi:10.1162/1064546053278946](https://doi.org/10.1162/1064546053278946)
- [125] G. M. Edelman, "Vaster than the Sky. The Gift of Consciousness," Yale University Press, New Haven/London, 2004.
- [126] G. M. Edelman, "Building a Picture of the Brain," *Daedalus*, Vol. 127, 1998, pp. 37-69.
- [127] A. Berthoz, "La Direction du Mouvement," Odile Jacob, Paris, 1997.
- [128] J. P. Tassin, "Le Rêve Naît du Sommeil," *Pour la Science*, Vol. 28, 1995, pp. 22-23.
- [129] M. Merleau-Ponty, "Phenomenology of Perception," Humanities Press, New York, 1962.