

A CONCISE COMPARISON OF SELECTED STUDIES OF CONSCIOUSNESS

Petros A. M. Gelepithis

Kingston University, Kingston upon Thames, England

Abstract

The literature on consciousness is huge and increasing with an accelerating pace. Moreover, the contributing scientists come from disciplines as varied as philosophy, physics, psychology, neurobiology, mathematics, computer science, and medicine. These two facts have created a multitude of viewpoints and have left little time for considering someone else's ideas, let alone trying to compare and eventually synthesize at least some of the various conceptions. This paper is a first step in that direction. It aims to put some order in the field and initiate as wide a comparative discussion of consciousness studies as possible.

To this end, we specify a class, L , of topics that, *presently*, constitute the major issues in the study of consciousness. Naturally, the complete class of issues, L_t , that an ideal theory of consciousness should, eventually, successfully address is not known. Nevertheless, L is both a useful approximation to L_t , and can be used as a reference framework for comparisons. Subsequently, we compare, in table form, eleven consciousness studies in terms of their stance to: (i) the nature of consciousness; and (ii) the issues constituting L . The eleven theories are coded under the following subheadings: Philosophical, Psychological, Computational, Neurobiological, and Biophysical. The paper concludes with some preliminary remarks and a statement of intended further work.

1. The Current Scope of Consciousness

“the difficulty, it seems, is that there is no agreement in the scientific community on what it is that we are actually studying when we do research on consciousness.” (Revonsuo et al. 1994, p. 20).

The scope of a discipline may be specified either definitionally or on the basis of issues studied. Definitions presuppose either a mature discipline incorporating a generally accepted, albeit approximate theory, or mathematical treatment of a subject matter. As Sommerhoff (1996 p. 139) has remarked:

“A precise definition of the word [‘consciousness’], of course, can only be the end-point of a theory of consciousness, just as the concepts of work and energy found a precise definition only as part of a theory of mechanics.”

The field of consciousness studies has quite some way to go to achieve such a maturity. So, it does not come as a surprise that current definitions of consciousness are inadequate as a sound basis for comparing competing theories of consciousness. Such definitions tend to specify a few aspects of consciousness and quite often they make arbitrary ontological or methodological assumptions.

The aim of this section, is to specify a class of issues on the basis of which a comparison of the diverse conceptions of consciousness will be made in the next one. In addition, such a class of issues can be used as an approximation to the complete class of issues that an ideal theory of consciousness should, eventually, successfully address. I will use the symbol L_t as an abbreviation for such a complete class of issues. It may be worth mentioning that L_t can not constitute a fixed, invariant class. Since human consciousness is a human cognitive phenomenon, its characteristics may well change with time. If not, that would mean we have come across a (new) cognitive invariant (see, e.g., Simon 1990; Gelepithis 1992). That would be interesting but I think improbable and it would definitely take us outside our current remit. We shall nevertheless suggest the invariant or – exclusive-or that is – time-dependent nature of consciousness as one of the issues in L_t .

Philosophers and scientists alike have tried to determine what needs to be explained. That is, to approach ‘consciousness’ through the identification of the major problems, phenomena, and questions assumed to fall within its scope.¹ Chalmers (1995) has been acknowledged to have made the distinction between two types of problems. First, ‘easy’ problems concerned with the finding of neuronal mechanisms underpinning perceptual and cognitive functions. For example, the difference between waking and sleeping, or how do anaesthetics work. Second, the ‘hard’ problems of trying to explain what are thoughts and feelings; how a neuronal system could give rise to the phenomenal experience of consciousness; how are we conscious of our perceptions.

Johnson-Laird (1987) specifies four major problems of consciousness. First, “to account for the distinction between what one can and cannot be aware of [what is going on in the world].” Second, “to draw a similar distinction [to that of awareness] between what one can and cannot consciously control.” Third, the problem of self-awareness. Fourth, the nature of intentionality (specified in contrast to goal-directedness).

Searle (1994) distinguishes eight features:

Subjectivity.

Unity (this is Searle’s term for the binding problem, namely, what are the mechanisms which bind the differentially processed elements of the visual world?).

Intentionality (this alleged feature is not really defined; only its general philosophical sense is given).

Field of consciousness (i.e., the class of all things which one is able to shift his/her attention to at any moment, for example feelings or thoughts).

Gestalt structure of conscious experience.

Familiarity (i.e., the capability of assimilating one's own experiences into a set of categories which are more or less familiar to oneself).

Mood.

Boundary conditions (i.e., "conscious states are situated and they are experienced as situated even though the details of the situation need not be part of the content of the conscious states." Searle 1994, p. 101).

Others (e.g., Edelman 1992; Gray 1995) have specified certain requirements that a theory of consciousness should fulfil. More recently, Edelman and Tononi (2000) have pointed out explicitly that neither introspection nor studies of the brain proper alone can provide a full answer to the riddle of consciousness. This is a point raising a number of significant questions and we shall return to consider it briefly in our discussion in section three.

All the above contributions to the nature of consciousness are useful approximations to . Below we try to integrate these contributions by creating a list, L , of the various topics that, presently, constitute the major issues in the study of consciousness. L is itself an approximation to L_t but a less incomplete one. Although obvious to me, it may be worth mentioning that L is not complete and that we humans may never become able to develop L to be identical to L_t . It nevertheless serves the good purpose that is the first such list available. No doubt L will be modified, extended or shrink in the future. This is a challenge for all of us and that is how science progresses.

L currently consists of twelve issues Q_1 - Q_{12} . For each issue we provide a brief comment intended to sketch some aspects of that particular issue for the multi-disciplinary readership of this journal. Finally, I would like to make explicit the fact that the 12 issues do not partition L and most of them are actually related to, or overlap with, each other. In particular, Q_3 cut across all members of L . Q_{11} is a special case of Q_9 ; we consider it separately both because of its significance and of the fact that theories have been developed for the one but not for the other. Q_{10} is wide-open and hence vague. Its reference to knowledge and memory, for example, may lead someone to investigate a host of, if not all conceivable, problems like meaning, communication, and learning. Still, it would be a mistake either to exclude Q_{10} from L or to put it on a procrustean bed to bring it down to one's delimited conceptions. At this stage, at least, we need to be explorers rather than judges of appropriate scope. Of course one may reasonably question the strategy of including overlapping issues and she may rightly claim that the specification of unequivocal categories to constitute L should be a challenge for one to aim for. My response is that at the current stage of development of consciousness studies, such an attempt would have created artificial and possibly ill advised boundaries. It would also have meant that the contributions or omissions of certain theories would have been less clearly seen. I think a partitioned L is a task to be carried out at a more mature state in the study of consciousness.

*Q*₁: Is it possible to incorporate consciousness into science?

There are five possible viewpoints:

Yes; consciousness can be incorporated within the currently known laws of nature.

Yes; but, consciousness requires an extension of the currently known laws of nature.

No; consciousness is a mystery beyond human science (e.g., McGinn 1989, 1995)².

No; consciousness is not a real phenomenon and should, thus, be eliminated from science.

No; consciousness denotes a hotch-potch of mental phenomena and as such no theory can be developed.

*Q*₂: What is consciousness?

In particular, what is a 'conscious' process or a 'conscious' state?³ In other words how could one decide whether a given mental event is conscious? Could a robot be conscious? Could there be qualitatively different classes of consciousness, for example, human, machine, chimp, or extraterrestrial consciousness?

*Q*₃: The problem of qualia or the subjective - objective issue.

The core issue in consciousness studies; no theory has so far addressed it even approximately adequately. Perhaps all workers in consciousness studies, with the exception of those taking the fourth stance on *Q*₁, consider it to be the hardest. For a classical statement the reader is referred to Nagel (1974*1991). A fundamentally special case of this issue is the debate on the adequacy conditions of a 'scientific' theory of consciousness. A minefield.

*Q*₄: The integrative and attentional nature of consciousness.

This is not a very often cited issue and it is usually expressed as the binding and cocktail-party problems respectively. I decided to put these two problems as two aspects of a single issue because they are closely related to each other and therefore study of the one can provide illuminating insights for the other. For example, to integrate, one should be able to appropriately distinguish and the latter requires considerable attentional power. The mental mechanisms of integration should, therefore, be related to those of making distinctions. Recent studies on this issue include Bogen (1995), Crick (1994) Dennett (1991), Dennett and Kinsbourne (1992), Gray (1995).

*Q*₅: Is consciousness an invariant?

If yes, what are the common characteristics in the consciousness of homo habilis and homo sapiens sapiens? If not, in what respects the consciousness of our ancestors differs from ours? And in what respects the homo sapiens sapiens consciousness could or will change?

*Q*₆: Are the words 'consciousness', 'awareness' and 'experience' absolute synonyms?

A lot of workers use them interchangeably, others not. To my knowledge, no-one so far has tried either to spell out their differences or to show the equivalence of their referents. Is, for example, self-consciousness reducible to any or a combination of the sensory, motor, emotional, and cognitive aspects of awareness? Is it solely related to direct experience? If yes, are animals self-conscious?

*Q*₇: The issue of altered states of consciousness.

Altered states of consciousness can not be defined. They include states like daydreaming, experiences of ecstasy, hallucinations, and hypnosis. The following characteristics seem to be common along the vast majority of altered states of consciousness: distorted time perception; diminished sense of one's self; and decrease in inhibitions. Hypnosis (e.g., Hilgard's (1977) hidden observer).

*Q*₈: Account for the distinction between conscious (Cs) and unconscious (Ucs) processes, e.g.,:

- (i) How do Ucs events become Cs?
- (ii) Why do we have both accessible and inaccessible Ucs?
- (iii) The issue of Cs control (Johnson-Laird 1987).
- (iv) What are the prerequisites for conscious processing? Would, for example, focal attention be adequate for it?

*Q*₉: What is the relation between brain on the one hand and consciousness and the unconscious on the other?

As Searle (1994, p. 93) remarks, addressing successfully even a subset of this issue, like "How exactly do neurobiological processes in the brain cause consciousness?", may well be the most important scientific discovery of the present era. We know that certain modifications in the structure or chemistry of the brain can affect consciousness. For example, the blindsight phenomenon (Weiskrantz 1986) demonstrates that visual information can produce motor behaviour on behalf of a person without having produced a corresponding conscious experience of that visual stimulus (received only by her primitive visual system). More generally, disorders in a number of modalities have led to loss of (access to) consciousness but not to a loss of overall knowledge corresponding to the particular damaged modality. Relevant work includes: Landis et al. (1980) for dyslexia; Bauer (1984), De Haan et al. (1987) for prosopagnosia; Milberg et al. (1987) for aphasia. For several examples on dissociable aspects of consciousness see Young (1996)⁴.

*Q*₁₀: What is the relationship between consciousness and memory?

A huge topic with two major approaches to the nature of memory. The majority viewpoint considers memory representational, the minority view argues that memory is nonrepresentational. Naturally, the issue of representation and the associated issues concerning the notions of symbol, meaning, and signal have also to be considered.

Q_{11} : Does consciousness have causal powers?

This issue is a special case of Q_9 but we consider it separately both because of its significance and the fact that theories have been developed for the one but not for the other. The most widely accepted body of evidence for the efficacy of mental states is the placebo effect (e.g., Wall 1996). A potentially quite useful methodological question is whether there is a single mechanism for consciousness or many?

2. Comparison of Selected Studies of Consciousness

Two points are in order here in order to clarify and justify the heading and content of this section. First, in this paper, a theory is perceived as a body of knowledge enabling an appropriate user to draw explanations and predictions about its subject matter and to control existing and design new systems within its boundaries (cf. Newell 1990). In this sense, we have not yet a theory of consciousness proposed. What we do have are explanatory theories, predictive theories, and computational models of certain aspects of consciousness. I will consequently use the term studies to refer to any of these terms and only use 'theory' in one of the two qualified senses specified above. After all, assuming membership of L_t is not constant, a theory of consciousness will need to be modified to accommodate the changing membership of L_t . If, or when, that becomes the case the term *theory* in the sense introduced at the beginning of this paragraph, will be used in the sense we use it for classical or quantum mechanics.

The second point concerns the studies we selected for review in this paper. The literature on consciousness is huge and increasing rapidly. A comprehensive review would require a monograph. Within the scope of a journal paper one has to be selective. Justification of one's selection on the basis of familiarity or taste alone would sound inadequate to me. To combine comprehensiveness with limited length, I decided to include in this paper the set of studies that are most appealing and familiar to me and follow it up with a second paper on the currently less familiar studies. As a prelude to this further work the next section cites the additional studies to be compared.

Our comparative review is presented in Table-1. The following subsections have been written with the sole purpose of clarifying the brief descriptions, or just terms within them, quoted or summarised in Table-1. In some cases, with the same objective in mind, brief supportive comments for the critical points summarised in Table-1 are given. The reader should also keep in mind that my classification of the compared studies under the headings of philosophical, psychological, computational, neurobiological, and biophysical indicate a disciplinary emphasis rather than an exclusive dependence on a particular approach or a set of findings.

Table-1: Comparison of selected studies of consciousness

(clarifications and precise quotation references in the relevant subsections following this table).

Consciousness Characterised As...	Critical Points
Baars 1994. Not specified. Claim: reflects a basic architectural aspect of the nervous system, functionally equivalent to a global workspace in a parallel and distributed set of neural processors."	A computational model concerning Q ₉ , and certain rather unrelated cognitive science topics. No contributions on Q ₂₋₈ & Q _{10, 11} . In summary, a theory of cognition rather than a theory of consciousness.
Chalmers 1995. Not specified. Claim: A theory of consciousness should take experience as a fundamental and irreducible notion.	An interesting claim on Q ₁ , and some food for thought in the form of his double-aspect principle for information. No contributions on Q ₂₋₁₁ .
Crick 1994. Not specified. A necessary condition for C is the production of significant reverberations by certain circuits in certain areas of the brain.	A "plausible model" for Q ₂ , Q ₄ , and aspects of Q ₇ . Hints on Q ₉ , accepts importance of Q ₁₀ . No contributions on the rest.
Eccles 1994. Not specified. It emerges as a consequence of the existence of psychons.	A 'heavy' model of Q ₉ , and Q ₁₁ , with a strong dualist claim concerning Q ₁ , and a forceful claim for the self-awareness aspects of Q ₆ . No contributions on the rest.
Edelman 1992. A "bootstrapping process" between V-CM and perceptual categorisation gives rise to PC.	An evolutionary model of Q ₂ , and Q ₉ , with a clear materialist claim concerning Q ₁ . Hints on Q ₆ , Q ₇ , Q ₈ , and Q ₁₀ . No contributions on the rest.
Freud 1938/†1940. Is only a quality of what is psychical, and an inconstant quality at that - one that is far oftener absent than present.	An extensive theory of major aspects of Q ₇ , and Q ₈ . Theoretical consideration of Q ₂ , Q ₁₀ , and Q ₁₁ . No contributions on the rest.
Libet 1994. A conscious mental field (CMF) that is a fundamental phenomenon irreducible to and not explainable by any known physical process. Such a field would be characterised by a unitary subjective experience and the ability to alter neuronal function.	A speculative, testable theory based on his hypothetical notion of the CMF aiming to address Q ₃ , Q ₄ , and Q ₉ . Needs considerable tidying up.

<p>Hameroff 1994. "macroscopic quantum state emerging from a critical level of quantum coherence occurring in and around cytoskeletal microtubules throughout the brain."</p> <p>Penrose 1994. "some manifestation of this quantum-entangled internal cytoskeletal state and of its involvement in the interplay (OR) between quantum and classical levels of activity."</p>	<p>A controversial model of Q₉, addressing significant issues of Q₁, (e.g., non-computability, causality, and aspects of space-time geometry), and aspects of Q₈. No contributions on the rest.</p>
<p>Searle 1994. the set of subjective states of awareness which end when one finds him/herself in unconscious states like, sleep, coma, etc.</p>	<p>No theory developed. Accepts the significance of most of the Q₁₋₁₁. No mention of Q₅.</p>
<p>Sommerhoff 1996. Not specified. The contents of primary consciousness are the contents of the postulated IGR.</p>	<p>A representational model of consciousness casting some light on Q₄. No direct contributions to the rest of the issues in L.</p>

2.1. *Philosophical Theories of Consciousness: Chalmers, Searle*

Within philosophical theories, a fundamental controversy concerns the ontological status of consciousness. For example, identity theory (see, e.g., Feigl 1958; Place 1956) attempts to identify raw feels (qualia in recent jargon) with physical properties of the nervous system. On the other hand, consciousness is often considered as an irreducible characteristic of reality like mass or space-time (see, for example, Chalmers 1995; Nagel 1986; Searle 1992).

Chalmers takes this viewpoint a step further by attempting to develop a theory of consciousness on the suggestion that such a theory should take experience as fundamental. Now, given that Chalmers has proposed the distinction between soft and hard problems of consciousness, and championed the importance of the latter, one would have expected his proposal to at least attempt to deal with the hard problem of consciousness, namely, "the problem of experience." (Chalmers, 1995, p. 201). By taking experience as a fundamental the hard problem of consciousness is essentially, once more, swept under the carpet. As Chalmers (1995, p. 210) himself admits:

"Of course, by taking experience as fundamental, there is a sense in which this approach does not tell us why there is experience in the first place."

Furthermore, it is far from clear whether 'experience' is taken as a fundamental property, or as a fundamental entity.

His theoretical proposal is based on three psychophysical principles: (i) the principle of structural coherence; (ii) the principle of organisational invariance; and (iii) the principle of the double-aspect of information. We concentrate our attention to his third "basic" principle since the other two

"involve high-level notions such as 'awareness' and 'organization', and therefore lie at the wrong level to constitute the fundamental laws in a theory of consciousness. Nevertheless, they act as strong constraints." (ibid. pp. 215-216).

The proposed double-aspect principle of information reads:

"information (or at least some information) has two basic aspects, a physical aspect and a phenomenal aspect". (ibid. p. 216).

Here the key notion of 'information' is proposed to be understood, "more or less", in Shannon's sense. Unfortunately, (i) Shannon's theoretical notion of information is admittedly, by Shannon himself, inadequate for capturing any of the semantic or pragmatic aspects of 'information' (Shannon and Weaver 1949); and, (ii) Shannon's wider views on information have not been used. In addition, Chalmers double-aspect principle of information lacks any substantive support; as he himself remarks:

"Of course, the double-aspect principle is extremely speculative and is also under-determined, leaving a number of key questions unanswered." (Chalmers, 1995, p. 217).

In summary, the proposed elements for a theory of consciousness provide some food for thought but they are far from constituting a theory of consciousness.

Searle (1994) defines consciousness as an exemplary complement to unconsciousness with awareness as a primitive defining term. Specifically, he defines consciousness as the set of subjective states of awareness which end when one finds him/herself in unconscious states like, sleep, coma, etc. He then goes on to specify the relation of consciousness to brain processes as a causal relation where neurobiological processes of a lower level play the role of the causing agent and conscious states are brain features of a higher level. The functional elements of the lower level neurobiological processes are assumed, in accordance with the prevailing view, to be neurons and synapses. On this basis he proceeds to introduce some of the key features of consciousness that an empirical theory of the brain should be able to explain. It is worth noting that, methodologically, Searle argues against both behaviourism and computationalism as viable approaches to the study of consciousness.

2.2. *Psychological Theories of Consciousness, Freud*

I was repeatedly amazed to read papers on 'consciousness' which so blatantly displayed either their ignorance of fundamental literature or their indifference for acknowledging it. The ignored or unacknowledged work is, of course, that of Freud. In summary, Freud's work was based on the belief that the unconscious constitutes the essence of psychological phenomena, and consciousness is our key methodological tool for uncovering the secrets and, possibly, laws of mental life. The following excerpts support and illuminate this summary remark and show that some of Freud's remarks are still state of the art, for example:

"Experience goes on to show that a psychical element (for instance, an idea) is not as a rule conscious for a protracted length of time. On the contrary, a state of consciousness is characteristically very transitory; an idea that is conscious now is no longer so a moment later, although it can become so again under certain conditions that are easily brought about. In the interval the idea was - we do not know what." (Freud, 1923*1991 p. 440).

And as Strachey⁵ remarks 'idea' in the above has been used to translate the German word 'Vorstellung' which covers three English terms: 'idea', 'image', and 'presentation'. Furthermore,

"No; being conscious cannot be the essence of what is psychical. It is only a quality of what is psychical, and an inconstant quality at that - one that is far oftener absent than present. The psychical, whatever its nature may be, is in itself unconscious and probably similar in kind to all the other natural processes of which we have obtained knowledge." (Freud, 1938/†1940*1991, p. 186).

Nevertheless,

"But none of this implies that the quality of being conscious has lost its importance for us. It remains the one light which illuminates our path and leads us through the darkness of mental life. In consequence of the special character of our discoveries, our scientific work in psychology will consist in translating unconscious processes into conscious ones, and thus filling in the gap in conscious perception..." (Freud, 1938/†1940*1991, p. 189).

However revealing the above excerpts may well be, they fall far short of even indicating the richness of Freud's work. The reader is strongly advised to consider at least some of Freud's relevant work; Anna Freud's (1991) selection is an excellent starting point.

2.3. Computational Theories of Consciousness: Baars

We confined our attention to Baars because, so far his work is clearly the best developed computational theory of consciousness. Earlier and much less developed information processing models of consciousness include those of Johnson-Laird (1983), and Schacter (1988). Michie (1994) summarises some of the best-known attempts to build artificial intelligence where consciousness is delegated to a secondary, if any, role and concludes that the disappointing results should prompt interest in reinstating consciousness “as the central arena for the mental solution of problems.” Unfortunately, Michie provides no reason to support his idea and makes no suggestion as to how consciousness may contribute to the mental solution of problems in the field.

Baars’ theory is based on the assumption that “the nervous system can be treated as a collection of unconscious specialized processors” which are recursively organised (Baars 1994). Its bare bones consist of 3 constructs:

- (i) a set of specialised semiautonomous processors
- (ii) a memory accessible by input processors, called global workspace (GW); and
- (iii) contexts, i.e., “unconscious knowledge that constrains conscious contents without itself being conscious.” (ibid.)

These three constructs are supplanted by the two processing principles of the GW theory, namely:

- competition between GW inputs; and
- global dissemination of output.

According to Baars (1994, p. 158) these 2 processing principles are considered helpful in searching for the neural correlates of consciousness. Unfortunately, both this and his other claim that consciousness’ “primary role is to recruit unconscious resources, to supervise, facilitate error-correction, and the like” have not been substantiated.

2.4. Neurobiological Theories of Consciousness:

Crick, Edelman, Libet, Sommerhoff

Crick’s (1994) proposal for a model of consciousness can be summarised under the following three interrelated hypotheses.

- H_1 : Consciousness is associated with certain neural activities. “This activity is largely in the lower cortical layers” V and VI, and “expresses the local (transient) results of “computation” taking place mainly in other cortical layers”. That is, cortical laminae I-IV.
- H_2 : Pyramidal cells in laminar V is the most likely type of cortical neurons that can express Consciousness. Activity in V and VI will not reach Consciousness unless sustained by some form of very short term memory.
- H_3 : The above sustenance may need a reverberatory circuit (RC) from VI → Thalamus → IV and VI. If a RC is lacking or if IV is too small, “it may not be possible to sustain these reverberations.”

A set of cortical areas is strongly connected to just one small region of the Thalamus (T_j). T_j “coordinates the activities of its associated cortical areas by synchronizing their firing”. T is intimately involved in attentional mechanisms.

Special binding takes the form of coordinated firing, often with rhythms in the 40-Hertz range.

Finally, Crick (1994, p. 252) writes:

“The regions involved in consciousness can influence, not necessarily directly, parts of the voluntary motor system. (There may be unconscious operations - “thoughts” - in between).

To repeat: Consciousness depends crucially on thalamic connections with the cortex. It exists only if certain cortical areas have reverberatory circuits (involving cortical layers 4 and 6) that project strongly enough to produce significant reverberations.”

So, as stated in Table-1, for Crick a necessary condition for consciousness is the production of significant reverberations.

Edelman’s (1989, 1992) early⁶ theory of consciousness is heavily evolutionary and is based upon a distinction, which he himself believes to be fundamental, between primary consciousness (PC) and higher-order consciousness (HOC). The former “is the state of being mentally aware of things in the world” while at the same time “it is not accompanied by any sense of a person with a past and future.” The latter “embodies a model of the personal, and of the past and the future as well as the present.”

Edelman’s four requirements for an evolutionary theory of consciousness may be summarised as follows: (i) it must propose explicit neural models that explain how consciousness arises; (ii) it must explain how consciousness emerges during evolution and development; (iii) it must connect consciousness to other mental matters such as concept formation, memory, and language; and (iv) it must describe stringent test for the models it proposes in terms of neurobiological facts (Edelman 1992, p. 113).

In the good old fashioned scientific way Edelman goes on to spell out the three hypotheses underlying his proposed theory:

H_1 : “the description of the world by modern physics is an adequate but not completely sufficient basis for a theory of consciousness.” (ibid.).

H_2 : consciousness arose as a phenotypic property at some point in the evolution of species” and, as a result of that, its acquisition “either conferred evolutionary fitness directly on the individuals having it, or provided a basis for other traits that enhanced fitness.” (ibid.).

H_3 : qualia exist in human beings other than one’s own self.

On this basis Edelman specifies three functions which, jointly, are necessary and sufficient for the appearance of PC:

f_1 : “the development of the cortical system in such a way that when conceptual functions appeared they could be linked strongly to the limbic system, extending already existing capacities to carry out learning.”⁷ (Edelman 1992, p. 119).

f_2 : the development of a new kind of memory, called value-category, that is based

on the link(s) between the limbic system and the conceptual functions. The value-category memory “is able to categorize responses in the different brain systems that carry out perceptual categorization and it does this according to the demands of limbic-brain stem values systems.” (ibid.).

f_3 : development of special reentrant circuits, in each modality, which allow “for continual reentrant signaling between the value-category memory and the ongoing global mappings that are concerned with perceptual categorization in real time.” (ibid.).

According to Edelman (1992, p. 119), it is this interaction (assumed to be a “bootstrapping process”) between V-CM⁸ and perceptual categorization that gives rise to PC. The question though is how does such a bootstrapping process exactly do that? His answer is because it “takes place in all sensory modalities in parallel and simultaneously, thus allowing for the construction of a complex scene”.

Although Libet’s work dates back to the mid sixties, his most interesting contribution with respect to a theory of consciousness is much more recent (Libet 1994; also Libet 1993a, 1993b, 1996a for excellent related experimental work⁹). To clarify the points summarised in Table-1, I would like to make three remarks. First, the conscious mental field (CMF) is assumed to be characterised by three properties:

- a) it is not describable in terms of any known physical theory.
- b) it unifies (or is a unified) subjective experience.
- c) it can cause modifications to neuronal processes.

Second, the proposed testability of CMF is very difficult to be carried out (see, e.g., Libet 1996b). Finally, it is not clear (Libet 1994 p. 120; 1996, p. 35) whether the integrative nature of the hypothetical CMF is confined to conscious phenomena or purports to cover the full gamut of mental events.

Sommerhoff (1996) “offers an account of consciousness as a biological process”. The account is couched in functional and causal terms and is based on his notion of an Integrated Global Representation (IGR). In essence the proposed IGR is a system of structures combining representations of actual or imagined objects, events, or situations. Naturally, the system of structures itself is an internal representation. This representational model of consciousness leads to the interesting notion of a state of expectancy characterised by two components: a state of readiness for an event X; and a state in which failure of X to occur elicits a tendency to initiate an orienting reaction. This notion of expectancy is used to claim that the integration of the various features of an object in the brain –the notorious binding problem- is accomplished “by the corresponding what-leads-to-what expectancies.” (ibid. p 1560. I think that this claim, very interesting as it is, has to be argued much more fully.

2.5. Biophysical Theories of Consciousness: Penrose, Hameroff, Eccles

Penrose (1994) believes that consciousness can not be explained within a computational framework and that, furthermore, a non-computational model for

consciousness requires new physical laws. In summary his suggestion reads:

“consciousness would be some manifestation of this quantum-entangled internal cytoskeletal state and of its involvement in the interplay (OR) between quantum and classical levels of activity.” (Penrose 1994, p. 376)¹⁰.

Let us see now more closely how the three key notions, namely consciousness, cytoskeletal state, and the OR physical theory are related to each other and what assumptions and experimental findings are used to support them.

First, the cytoskeleton consists of various protein-like molecules like microtubules (i.e., hollow cylinders with outside diameter of 25nm and inside one of 14nm)¹¹. An appropriate class of microtubules, across the brain, are assumed to be capable of sustaining quantum-coherent states or at least to provide “sufficient quantum entanglement” so that a classical description is inappropriate.

Second, it is assumed that the derived quantum entanglement is necessary for consciousness.

Third, the quantum/classical interface described by the missing OR theory influences the brain’s behaviour “through the cytoskeletal control of synaptic connections.” (Penrose 1994, p. 371).

It should be noted that quantum coherence¹² is assumed to occur among tubulins in brain microtubules¹³, and that it continues until a quantum gravitational threshold is reached. It is furthermore assumed that microtubule-associated proteins are responsible for the wave function collapse¹⁴. On the question of how did consciousness arise Penrose (1994*1995) and Hameroff (1994) differ. The former argues for non-computational physical laws as the necessary background within which the brain’s complex organisation found suitable ground for the development of consciousness. Hameroff on the other hand argues that consciousness should be viewed as an emergent property of physical systems.

Eccles’ (1994) *How the Self Controls Its Brain* brings together the results of his life-long study of the mind-brain problem. It culminates in proposing a hypothetical mechanism

“the scientific hypothesis of how the self may control its brain without infringing the conservation laws.” (ibid. p. 181).

Specifically, the hypothesis reads:

“the mental intention (the volition) becomes neurally effective by momentarily increasing the probability of exocytosis in selected cortical areas such as the supplementary motor area neurons”. (Beck & Eccles, 1992*1994, p. 160).

Eccles’ hypothesis is based upon a quantum mechanical description of bouton exocytosis and the experimental finding that exocytosis is an all-or-nothing event occurring with probabilities of about 0.2-0.3. The quantum mechanical model of exocytosis (argued to be the key mechanism of neocortical activity, Beck & Eccles, 1994)

is essentially based on the assumption that exocytosis is governed by a trigger mechanism which can, furthermore, be modelled by the motion of a quasi-particle. This mind-brain interaction is explained in quantum mechanical terms.

3. Preliminary Conclusions, Remarks, and Further Work

Based on the studies compared so far, the following preliminary conclusions can be drawn. First, none of the studies reviewed so far has tackled either Q_3 , except in a speculative mode, or Q_5 ! Second, no study has addressed, exhaustively and in depth, more than 2 of the 11 issues. This point is made only in order to make crystal clear the tremendous amount of work that is still required. This leads naturally to our third remark, namely, that no study has attempted to integrate the work that has been done so far. To a large extent this is due to the variety of the approaches involved, not all of which are necessarily compatible with each other. Nevertheless, although integration may well be impossible, selective synthesis should be feasible and I believe necessary on the road towards a theory of consciousness. I furthermore believe that such a synthesis should be attempted within the overall theoretical framework required, and still, to be developed for the study of mind and brain.

We intend to write a sequel paper reviewing the major remaining studies of consciousness and outlining a preliminary synthesis along the lines sketched above. Our intention is to review:

- Baddeley's (1993) idea of consciousness as "one of the functions of the central executive component of working memory".
- Block's (1995) scope of consciousness and the function of phenomenal consciousness.
- Churchland's (1995) recurrent-network conception of consciousness.
- Dennett's (1991*1992) multiple drafts model.
- Dretske's (1995) representational 'theory' of consciousness.
- Gray's (1993, 1995) model of anxiety and schizophrenia and his subicular comparator hypothesis.
- Natsoulas' (1978, 1999a, 1999b) work on the nature of consciousness.
- Rosenthal's (1986) reductive study of consciousness.
- Taylor's (1999) model of mind and relational consciousness.
- Velmans (2000) *Understanding Consciousness*.
- Weiskrantz's (1997) proposal on the nature and character of consciousness.

Naturally, the eventual scope of our review will depend on the progress in consciousness studies.

Acknowledgements: I would like to thank Gerhard Dalenoort and an anonymous referee for their useful comments on an earlier version of this paper. Their remarks contributed to a better paper.

References

- Baars, B. J. (1994), A Global Workspace Theory of Conscious Experience. In: Revonsuo, A. and Kamppinen, M. (Eds.), *Consciousness in Philosophy and Cognitive Neuroscience*, Lawrence Erlbaum Associates.
- Baddeley, A. (1993), Working Memory and Conscious Awareness. In Collins, A. F., Gathercole, S. E., Conway, M. A., and Morris, P. E. (Eds.), *Theories of Memory*, Lawrence Erlbaum Associates.
- Bauer, R. M. (1984), Autonomic recognition of names and faces in prosopagnosia: A neuropsychological application of the guilty knowledge test, *Neuropsychologia*, **22**, 457-469.
- Beck, F., and Eccles, J. C. (1994), Quantum Aspects of Brain Activity and the Role of Consciousness, in: Eccles, J. C., *How the Self Controls its Brain*, Springer-Verlag, Berlin, (All but the introduction of this chapter first appeared in *Proc. Nat. Acad. Sci.*, **89**, 11357-11361.
- Block, N. (1995), On a Confusion on a Function of Consciousness, *Brain and Behavioral Sciences*, **18**, 227-247.
- Bogen, J. E. (1995), On the neurophysiology of consciousness: I. An overview, *Consciousness and Cognition*, **4**, 52-62.
- Chalmers, D. (1995), Facing up to the problem of consciousness, *Journal of Consciousness Studies*, **2**, 200-219.
- Churchland, P. M. (1995), *The Engine of Reason, the Seat of the Soul: A Philosophical Journey into the Brain*, MIT Press, Cambridge (Mass.).
- Crick, F. H. C. (1994), *The Astonishing Hypothesis: The Scientific Search for the Soul*, Touchstone Books, London.
- De Haan, E. H. E., Young, A., and Newcombe, F. (1987), Face recognition without awareness, *Cognitive Neuropsychology* **4**, 385-415.
- Dennett, D. C. (1991*1992), *Consciousness explained*, Allen Lane, The Penguin Press.
- Dennett, D. C., and Kinsbourne, M. (1992), Time and the observer: The where and when of consciousness in the brain, *Brain and Behavioral Sciences* **15**, 183-247 (including Open Peer Commentary),
- Dretske, F. (1995), *Naturalizing the Mind*, MIT Press, Cambridge (Mass.).
- Eccles, J. C. (1994), *How the Self Controls its Brain*, Springer-Verlag, Berlin.
- Edelman, G. M. (1989), *The Remembered Present: A Biological Theory of Consciousness*, Basic Books, New York.

- Edelman, G. M. (1992), *Bright Air, Brilliant Fire: On the Matter of the Mind*, Basic Books, New York.
- Edelman, G. M., and Tononi, G. (2000), *Consciousness: How Matter Becomes Imagination*, Penguin Books.
- Feigl, H. (1958), *The 'mental' and the 'physical'- The essay and a Postscript*, University of Minnesota Press, Minneapolis.
- Freud, A. (1991), *The Essentials of Psycho-Analysis*, Penguin Books.
- Freud, S. (1923*1991), The Ego and the Id, in: Freud, A., *The Essentials of Psycho-Analysis*, Penguin Books.
- Freud, S. (1938*1991), Some Elementary Lessons, in: Psycho-Analysis, in Freud, A. *The Essentials of Psycho-Analysis*, Penguin Books.
- Gelepithis, P. A. M. (1992), True invariants: A response to Herbert Simon, *Proceedings of the 13th International Congress on Cybernetics*, Namur, Belgium, 24-28 August 1992, pp. 424-428.
- Gray, J. (1993), Consciousness, schizophrenia and scientific theory, in: Bock, G. R., and Marsh, J. (Eds.), *Experimental and Theoretical Studies of Consciousness*, (CIBA Foundation Symposium 174). John Wiley & Sons.
- Gray, J. (1995), The contents of consciousness: a neuropsychological conjecture. *Behavioral and Brain Sciences* 18, 659-722 (including commentaries).
- Grush, R., and Churchland, P. S. (1995), Gaps in Penrose's Toilings, *Journal of Consciousness Studies* 2, 10-29.
- Hameroff, S. R. (1994), Quantum coherence in microtubules: A neural basis for emergent consciousness, *Journal of Consciousness Studies* 1, 91-118.
- Hilgard, E. R. (1977), *Divided Consciousness: Multiple Controls in Human Thought and Action*, Wiley-Interscience.
- Johnson-Laird, P. (1983), *Mental Models*, Harvard University Press.
- Johnson-Laird, P. (1987), How Could Consciousness Arise from the Computations of the Brain?, in: Blakemore, C., and Greenfield, S. (Eds.), *Mindwaves*, Basil Blackwell, Oxford.
- Landis, T., Regard, M., and Serrant, A. (1980), Iconic reading in a case of alexia without agraphia caused by a brain tumour: A tachistoscopic study, *Brain and Language* 11 45-53.
- Libet, B. (1993a), *Neurophysiology of consciousness*, Birkhäuser, Boston.

- Libet, B. (1993b), The Neural time factor in conscious and unconscious events, in: Bock, G. R., and Marsh, J. (Eds.), *Experimental and Theoretical Studies of Consciousness*, (CIBA Foundation Symposium 174), John Wiley & Sons.
- Libet, B. (1994), A testable field theory of mind-brain interaction, *Journal of Consciousness Studies*, 1, 119-126.
- Libet, B. (1996a), Neural processes in the production of conscious experience, in: Velmans, M., (Ed.), *The Science of Consciousness: Psychological, Neuropsychological and Clinical Reviews*, Routledge, London.
- Libet, B. (1996b), Solutions to the hard problem of consciousness, *Journal of Consciousness Studies*, 3, 33-35.
- McGinn, C. (1989), Can we solve the mind-body problem? *Mind* 98, 349-366.
- McGinn, C. (1995), Consciousness and space, *Journal of Consciousness Studies* 2, 220-230.
- Michie, D. (1994), Consciousness as an engineering issue, part I, *Journal of Consciousness Studies* 1, 182-195.
- Milberg, W., Blumstein, S. E., and Dworetzky, B. (1987), Processing of lexical ambiguities in aphasia, *Brain and Language* 31, 138-150.
- Nagel, T. (1974*1991), What Is It Like to Be a Bat?, in: Rosenthal, D. M. (Ed.), *The Nature of Mind*, Oxford University Press.
- Nagel, T. (1986), *The View from Nowhere*, Oxford University Press.
- Natsoulas, T. (1978), Consciousness, *American Psychologist* 33, 906-914.
- Natsoulas, T. (1999a), The Concept of Consciousness: The General State Meaning, *Journal for the Theory of Social Behaviour* 29, 59-87.
- Natsoulas, T. (1999b), An Ecological and Phenomenological Perspective on Consciousness and Perception: Contact With the World at the Very Heart of the Being of Consciousness. *Review of General Psychology*, 3, 224-245.
- Newell, A. (1990), *Unified Theories of Cognition*, Harvard University Press.
- Niiniluoto, I. (1994), Scientific Realism and the Problem of Consciousness, in: Revonsuo, A. and Kampinnen, M. (Eds.), *Consciousness in philosophy and cognitive neuroscience*, Lawrence Erlbaum.
- Penrose, R. (1994*1995), *Shadows of the Mind*, Vintage, London.
- Penrose, R., and Hameroff, S. (1995), What 'gaps'? Reply to Grush and Churchland. *Journal of Consciousness Studies*, 2, 98-111.

- Place, U. T. (1956*1970), Is consciousness a brain process, in: Borst, C. V. (Ed.), *The mind-brain identity theory*, Macmillan, St Martins Press.
- Revonsuo, A., Kamppinen, M. and Sajama, S. (1994), General Introduction: The Riddle of Consciousness, in: Revonsuo, A. and Kampinen, M. (Eds.), *Consciousness in philosophy and cognitive neuroscience*, Lawrence Erlbaum Ass.
- Rosenthal, D. (1986), Two Concepts of Consciousness, *Philosophical Studies* 49, 329-359.
- Schacter, D. L. (1988), On the relation between memory and consciousness: Dissociable interactions and conscious experience, in: Roediger, H., and Craik, F., (Eds.), *Varieties of Memory and Consciousness*, Lawrence Erlbaum Ass.
- Searle, J. R. (1992), *The Rediscovery of the Mind*, MIT Press, Cambridge (Mass.).
- Searle, J. R. (1994), The Problem of Consciousness, in: Revonsuo, A. and Kampinen, M. (Eds.), *Consciousness in philosophy and cognitive neuroscience*, Lawrence Erlbaum. (Originally published in *Experimental and theoretical studies of consciousness*, Wiley, Chichester; CIBA Foundation Symposium 174).
- Shannon, C. E., and Weaver, W. (1949), *The Mathematical Theory of Communication*, University of Illinois Press.
- Sommerhoff, G. (1996), Consciousness Explained as an Internal Integrating System. *Journal of Consciousness Studies* 3, 139-157.
- Taylor, J.G. (1999), *The race for consciousness*, MIT-Press-Bradford, Cambridge, Mass.
- Velmans, M (2000), *Understanding Consciousness*, Routledge.
- Wall, P. D. (1996), The Placebo Effect, in: Velmans, M. (Ed.), *The Science of Consciousness: Psychological, Neuropsychological and Clinical Reviews*, Routledge, London.
- Weiskrantz, L. (1986), *Blindsight*, Clarendon Press, Oxford.
- Weiskrantz, L. (1997), *Consciousness lost and found: A neuropsychological exploration*, Oxford University Press.
- Young, A. W. (1996), Dissociable aspects of consciousness, in: Velmans, M. (Ed.), *The Science of Consciousness: Psychological, Neuropsychological and Clinical Reviews*, Routledge, London.

Notes (the page numbers between brackets refer to this article)

- 1 For some people consciousness is identified with mind (e.g., Niiniluoto 1994; Searle 1992). This is an exaggeration we do not share. Consciousness, although related, it is different from, say, emotion and thinking. Naturally, one may say that such an identification refers only to the nature of the mind or consciousness. In this case the claim is true but it is superfluous; it is against both simplicity and cognitive economy. (p. 374)
- 2 He presents it as a possibility rather than as a strong claim. (p. 376)
- 3 I will use the term 'conscious' phenomena (i.e., states, processes, and their combination) as an abbreviation whenever such usage creates no problem. (p. 376)
- 4 The reader must keep in mind that the area of brain-consciousness studies is vastly wider than the work referenced here indicates. (p. 377)
- 5 Editorial note taken from the Standard Edition of the Complete Psychological Works of Sigmund Freud and repeated *ibid.* p. 149. (p. 382)
- 6 His and Tononi's expanded views (Edelman and Tononi 2000) deserve, in our opinion, a full review paper that falls outside the scope of this comparison. (p. 384)
- 7 What f_1 amounts to is, essentially, a further requirement for an appropriate cortical system to be developed which seems to put the horses before the cart. Why did the cortical system have to be developed in such a way that when conceptual functions appeared they could be linked strongly to the limbic system? (p. 384)
- 8 Do the V-CM categorisations follow exclusively from the demands of the limbic system? No clear answer to this question is provided. (p. 385)
- 9 The reader is advised to consider carefully the discussion following Libet 1993b). (p. 385)
- 10 For a very critical review see Grush and Churchland (1995), for an equally vociferous reply Penrose and Hameroff (1995). (p. 386)
- 11 In general, each microtubule consists of 13 columns of tubulin dimmers (approximate dimensions 8x4x4 nm) which, in turn, consist of 2 parts (a-tubulin and b-tubulin) composed of approximately 450 aminoacids each. (p. 386)
- 12 That is, large numbers of particles co-operating in a single quantum state which remains unentangled with its environment. (Penrose 1994*1995, p. 351). (p. 386)
- 13 An appropriate class of microtubules, across the brain, are assumed to be capable of sustaining quantum-coherent states or at least to provide "sufficient quantum entanglement" so that a classical description is inappropriate. (p. 386)
- 14 The quantum/classical interface, described by the missing OR theory, influences the brain's behaviour "through the cytoskeletal control of synaptic connections." (p. 386)

Manuscript received: 4-7-2000, Final Version 22-1-2001

Address Author:

Kingston University, Dept. of Computer Science
Kingston upon Thames, KT1 2EE, England

~~E-mail: petros@kingston.ac.uk~~

petros.gelepithis@philosophy.ox.ac.uk
petros2004@phs.uoa.gr