Brain Processes and Phenomenal Consciousness

A New and Specific Hypothesis[1]

Hans Flohr
UNIVERSITY OF BREMEN

ABSTRACT. A hypothesis on the physiological conditions for the occurrence of phenomenal states is presented. It is suggested that the presence of phenomenal states depends on the rate at which neural assemblies are formed. Unconsciousness and various disturbances of phenomenal consciousness occur if the assembly formation rate is below a certain threshold level; if this level is surpassed, phenomenal states necessarily result. A critical production rate of neural assemblies is the necessary and sufficient condition for the occurrence of phenomenal states.

Phenomenal experience, qualia or raw feels are badly defined terms (Dennett, 1988) denoting the phenomenal, subjective, qualitative properties of experiences directly accessible to introspection. Experiencing pains, colors, sounds, smells, etc., are typical examples.[2] The relation of phenomenal states to states and processes in the brain is the subject of controversial debate. One standpoint (Bieri, 1987; Farrel, 1962; Jackson, 1982; Nagel, 1974, 1986) holds that phenomenal states cannot *in principle* be explained by neurophysiological facts. There are three main families of arguments against a physicalistic explanation of the phenomenal aspects of mentality.

1. The *explanatory gap* argument claims that statements on identity of brain states and phenomenal states are unintelligible and leave an explanatory gap. We have no conception of how an event or state in the brain could have phenomenal properties. A physicalistic description of brain states and processes fails to grasp the essence of what is a phenomenal state. It cannot provide answers to questions such as: 'How is it to be in pain?' (Bieri); 'What is it like to be and hear like a bat?' (Farrel); or 'How does it taste to eat cauliflower?' (Dennett). Physiological analysis provides no information on the subjective, experiential aspect of sensations. Even if the physicalistic description of brain events was more

THEORY & PSYCHOLOGY © 1991 SAGE. VOL. 1(2): 245–262

detailed and more accurate the situation would remain unchanged. Even if behavior could be completely explained in neurophysiological terms, qualia would still be left unexplained. Qualia are left out of the physicalistic story because any objective physicalistic view of consciousness necessarily involves abstracting from the personal, subjective view; or, as Thomas Nagel (1986) expressed it: 'We will not know exactly how scrambled eggs taste to a cockroach even if we develop a detailed phenomenology of the cockroach sense of taste'. Physics—according to this view—is incomplete, 'bound to leave undescribed the irreducibly subjective character of conscious mental processes whatever may be their intimate relation to the physiological operations of the brain' (Nagel, 1974).

2. The *knowledge* argument (Farrel) maintains that there is some knowledge in having subjective experiences that can be acquired only by having the relevant experience itself. Knowledge of what goes on in the brain when a person has a specific experience does not include the information the subject has when it is in a certain phenomenal state.

3. The *absent qualia* argument is directed against the thesis that qualia are functional states and could be type-individuated by their causal role. It claims that, in principle, any functionally specified structure can be realized in the absence of qualia.

The opposite view (Churchland, 1985; Dennett, 1988; Lewis, 1983; Loar, 1990; Lycan, 1987, 1990; van Gulick, 1990) maintains that our notions of phenomenal states can be reduced to notions of states and activities of the brain, or need to be replaced completely by the more appropriate neurophysiological terminology. Firstly, because there is a number of points at which the above anti-reductionistic arguments can be attacked (van Gulick, 1990). Secondly, because it is unquestionable that the existence and the properties of phenomenal states depend on the functioning of the brain. For example, a short period of oxygen lack causes unconsciousness, and this dependence on the energy metabolism of the brain is a property that phenomenal states share with a number of physiological states. With the latter they belong to a class of states that critically depend on the thermodynamic state of the open system which the brain is. They occur only if this system is maintained in a non-equilibrium state by a permanent supply of free energy. The physiologist can name further conditions for the occurrence of phenomenal states. He or she can generate them (e.g. by direct electrical stimulation of the brain), modify them (e.g. by psychotropic drugs), or prevent them temporarily or permanently (e.g. by anesthesia or by lesions of circumscript brain regions). It may be that, at present, our knowledge of the physiological conditions for the occurrence of phenomenal states is not sufficient. But, in principle, nothing stands in the way of a complete account of these conditions. And it may turn out that such a 'complete' account would

constitute the basis for a naturalistic explanation of the phenomenon of subjectivity. In other words, it may be premature to anticipate a priori that these arguments will survive all future developments of the neurosciences as serious obstacles to a physicalistic account of the phenomenal aspects of mentality.

Conscious and Unconscious Brains

For the moment I would like to leave this open and concern myself with the unquestionable relationship between phenomenal states and states and processes in the brain. First, I try to delineate the differences in the states and processes occurring in the brain when it is aware of sensations, and when it is not—as far as that is presently possible; i.e. I try to describe the 'outer' objective aspects of systems that have or do not have 'inner' subjective aspects. Second, I propose a hypothesis on *why* certain systems develop phenomenal states and others do not.

The physiologist has three possible ways of detecting differences between conscious and unconscious brains and thus deducing the conditions for consciousness. First, one can try to register correlations, parallels and coincidences between certain brain states and certain phenomenal states, i.e. try to describe the physiological state of the brain when the person to whom this brain belongs is in a particular phenomenal state, and compare this description with those obtained in the absence of that particular phenomenal state. I would like to call this approach the correlative strategy. When such attempts are performed on experimental subjects other than oneself, the existence of certain phenomenal states can be assumed only on the basis of the subject's verbal reports or other behavior. Feigl (1958) and Meehl (1966) have therefore proposed that such an investigation should be conducted as a self-experiment, i.e. as autocerebroscopy. Apart from the experimental difficulties that such an approach would present, there are other, more fundamental problems that have been dealt with elsewhere (Flohr, 1989). Second, the physiologist can intrude into the working brain and try to generate or modify phenomenal states or to prevent their occurrence. I would like to call this the intervention strategy. Included in this strategy are the clinical observations on functional deficits following brain lesions. Third and finally, the physiologist can combine both strategies and examine, for example, whether brain states that are regularly correlated with the occurrence of certain phenomenal states are capable of causing the same states if they are generated experimentally.

The findings of such experiments are usually ambiguous. A coherent theory can be formulated only on the basis of large numbers of individual observations. This task can be compared to the completion of a jigsaw

puzzle: one can be successful only if (a) one has a sufficient number of pieces and (b) an idea of what the final picture will look like.

The puzzle, as it appears at the moment, is very incomplete and fragmentary. Here, I give a very simplified and abbreviated account. The cerebral cortex receives two classes of inputs (Figure 1):

(a) *Specific* afferents from sense organs are relayed through specific thalamic nuclei before reaching the primary cortical projection areas. For instance, somatosensory information is transmitted via the nucleus

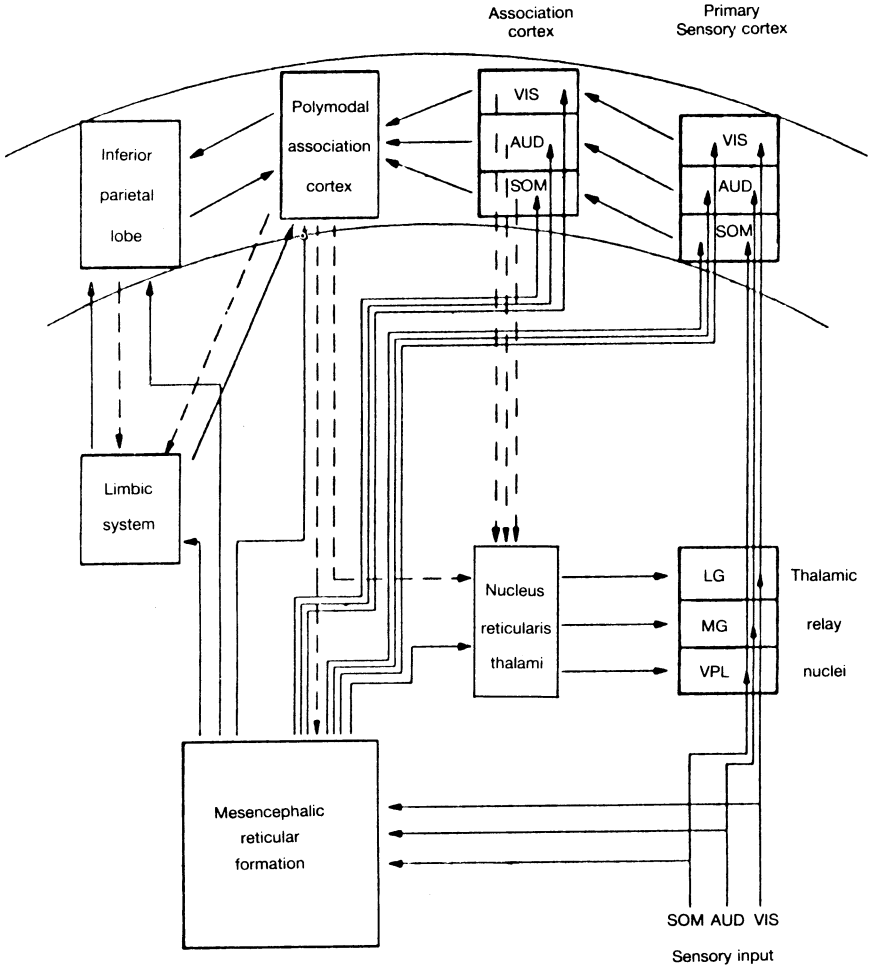

FIGURE 1. Specific and non-specific inputs to the cerebral cortex. SOM = somatosensory pathways; AUD = auditory pathways; VIS = visual pathways; LG = lateral geniculate; MG = medial geniculate; VPL = Nucleus ventralis posterolateralis.

ventralis posterolateralis to the postcentral gyrus; auditory information is transmitted through the medial geniculate nucleus to the supratemporal plane (Heschl's gyrus); visual afferents are transmitted through the lateral geniculate nucleus to area 17 of the occipital lobe.

(b) *Non-specific* afferents reach the cortex from the mesencephalic reticular formation. Reticular formation is the term applied to a network of neurons and nerve fibers extending from the caudal medulla to the diencephalon. A common feature of all these neurons is that they form a widely distributed network of synaptic contacts. In contrast to older accounts, which described the reticular formation as a mass of poorly organized neurons and fibers, it has now been established that it is organized into different nuclei with specific afferent and efferent connections. The reticular formation is co-innervated via axon collaterals from all sensory afferents ascending from the periphery to the cerebral cortex. Moreover, it receives afferents from the cortex, the basal ganglia and the cerebellum. The mesencephalic and diencephalic reticular formation projects rostrally to the cerebral cortex and influences the cortical processing of specific sensory afferents. These pathways comprise the ascending reticular activating system which is known to play an essential role in arousal, wakefulness and attention. There is a second site where the reticular formation influences the processing of primary afferents: the thalamic relay nuclei. The nucleus reticularis thalami, a thin sheet of neurons, surrounds the dorsal thalamus and inhibits the thalamic relay nuclei. Its control function is, in turn, affected by collaterals of thalamocortical pathways, by collaterals from cortico-thalamic projections and by inhibitory afferents from the mesencephalic reticular formation.

Within the cerebral cortex, primary auditory, somatosensory and visual fields each project to adjacent unimodal association areas which, in turn, project to contiguous secondary unimodal association fields. It is assumed that these successive areas are concerned with progressively higher-order abstractions within a modality. The unimodal association areas project to a number of polymodal sensory areas lying in the cingulate gyrus, parietal, temporal and frontal lobes. Their function is vaguely described as a crossmodal association and synthesis. The polymodal association areas project to the inferior parietal lobe which has been termed a 'supramodal' area. Polymodal and supramodal regions have connections to the limbic system; these connections provide the anatomical substrate by which motivational states influence cortical processing of sensory stimuli.

Lesions within this system lead to various disturbances of consciousness which are relatively well known. *Global* and *focal* disturbances of consciousness can be distinguished. An extremely simple summary would be as follows: bilateral lesions of an area of the mesencephalic reticular formation extending from the rostral pons to the diencephalon lead to coma, a global loss of consciousness. The subject cannot be aroused from

this form of unconsciousness, even by strong external stimulation (Saper & Plum, 1985). With focal disturbances of consciousness the person is awake, but has specific deficits in his conscious experience. Depending on the lesion, awareness of entire subsets of conscious states is lost. With lesions of the primary sensory areas (or their specific afferents), perceptions in this modality do *not* enter consciousness; they lead to cortical blindness, anesthesia or deafness. With lesions of the association cortex, more complex deficits result: the various forms of agnosias. Color agnosia is an example of the loss of a single category of phenomenal qualities. Discrete, unilateral lesions of the mesencephalic reticular formation, or of the ascending reticular activating system, lead to the (attentional) neglect syndrome (Heilman, Valenstein, & Watson, 1985). Neglect patients typically fail to report, respond, or orient to novel or meaningful stimuli presented to the side opposite the lesion. An example is hemi-inattention when sensory stimuli from one or more modalities on one side of the body are ignored. Patients with profound hemi-inattention may even fail to recognize that their extremities contralateral to the lesion are their own.

I would not have presented this simplified account if it was not sufficient to permit a preliminary conclusion and to develop a new, broader-based hypothesis. After what has been said, the difference between the conscious and unconscious brain consists in the manner in which the two streams of afferent signals are processed. Consciousness depends on the convergence of specific and unspecific afferent systems, both of which contain complex feedback loops and are interconnected at various levels. In this respect, the condition for consciousness is this convergence of afferent information. Consciousness arises 'somehow' as the consequence of the interactions between specific and unspecific afferents.

This relatively vague statement is the result of several decades of extensive research into the role of the brain stem in the generation and maintenance of consciousness. It fails to say anything concrete about the difference in the subsequent cortical processing of sensory data distinguishing conscious and unconscious brains. It says nothing about the specific nature of those brain processes which result from this convergence and whose occurrence is a precondition for the occurrence of consciousness. It is at this point that the *terra incognita* of the physiology of consciousness begins.

The Essential Difference

What we now need to know is twofold: (1) What happens beyond this point of confluence? (2) Which specific physiological states and processes result from this convergence of signals that could be constitutive for phenomenal experiences? I would like to divide the total of possible neural processes

resulting from the interaction between specific and unspecific afferents into two categories:

1. 'conventional' synaptic transmission, i.e. signal propagation in neural nets with rigid structure and given synaptic weights;
2. activity-dependent self-organization of neural nets, i.e. changes in synaptic connections which are induced by synaptic activation and dependent on the degree of activation.

We now know that the conventional view of brain function—electrical impulses spreading through a net of neurons connected by rigid synapses—has to be revised. The induction of postsynaptic potentials by neurotransmitters gating the opening of ion channels is only one mechanism by which neurons communicate. In addition, transmitters can activate receptors that are coupled to the production of intracellular second messengers, e.g. cyclic nucleotides, diaglycerol, inositol triphosphate and calcium ions, that initiate changes in neuronal excitability. These include lasting changes brought about by changes in gene expression or post-translational modification of neuronal proteins, as well as transient, dynamic changes in the strength of connections between neurons. Activity in neural nets is coupled with the reorganization of these nets.

In the following I present a speculative hypothesis claiming that the occurrence of phenomenal states depends critically on the presence of processes belonging to the second group. *It is suggested that the occurrence of phenomenal states depends on the rate at which activity-dependent synaptic changes occur and neural assemblies are formed.*

Assemblies

The concept of neural cell assemblies is central to this hypothesis. It was originally developed by D.O. Hebb (1949) and modified by C. von der Malsburg (1981). In *The Organization of Behavior*, Hebb proposed that

> . . .repeated stimulation of specific receptors will lead slowly to the formation of an 'assembly' of association area cells which can act briefly as a closed system after stimulation has ceased; this prolongs the time during which structural changes of learning occur. . . .

and he suggested that an assembly 'constitutes the simplest instance of a representative process (image or idea)'. The basis of this self-organization of neural nets is the presence of plastic synapses, now termed Hebb synapses. Hebb assumed that synapses on a neuron that are active while the neuron discharges will be strengthened, whereas inactive synapses will

be weakened. Synapses from differing inputs that are active at the same time on the same neuron will be reinforced and selected over others.

When applied to nets of spatially distributed groups of neurons with non-specific, random interconnections, these so-called Hebb rules lead to a relative stabilization and association of neurons firing in a correlated fashion. An *assembly* of preferentially connected, coherently active cells is formed. If *coincident* activity is induced in some neurons of such a net by a patterned input, an assembly will be formed because the synchronous activation selectively modulates the pathways connecting these neurons. The assembly detects and encodes the coherent properties of the stimulus pattern such that a representation of that pattern is generated. Once the assemblies have been formed, they would function as detectors of the same or similar input patterns expressing the detection of coherent features by co-ordinating their activities. It is easy to envisage that the output of such assemblies could in turn be used as input to other modifiable nets which then would reorganize their structure as a function of this input. Iteration of such processes would generate more and more abstract *metarepresentations*.

There has been much theoretical work done on such self-organizing associative nets (Bienenstock, 1983; Kohonen, 1983, 1984; Ritter & Kohonen, 1989; von der Malsburg, 1973). These nets are able to create topographical representations in which the relationship between elements of the stimulus pattern become encoded into spatio-temporal patterns of activities. Moreover, in such nets abstractions from primary sensory data at increasingly higher levels of generalization are formed and stored. Synaptic plasticity is the means by which neural nets detect and represent coincident properties and, hence, the nomological relations existing between elements of the input pattern.

Rapid Changes

The idea of von der Malsburg (1973, 1981, 1983; Wilshaw & von der Malsburg, 1976) was that the assembly formation is not a slow process but that rapid synaptic changes are possible allowing the formation of transient assemblies in much shorter periods of time (100–200 ms). Crick (1982) has termed these hypothetical synapses 'Malsburg synapses'. Assuming the existence of such synapses would be a rather radical departure from the usual assumption that in performing neural nets the synaptic weights are kept constant. The intriguing feature of this proposal is that the brain is conceived as something like a high-speed generator of temporary representations and metarepresentations implemented by the network's interconnections.

The existence of such rapid synaptic changes has received support

recently from a number of physiological observations. The essence of Hebb's postulate is that a synapse is strengthened if the presynaptic activation of the synapse is coincident with the firing of the postsynaptic neuron. This poses two difficulties:

(a) some mechanism of the synapse must respond to coincident pre- and postsynaptic activity and initiate temporary changes in the weights of the synapse;
(b) in the case of Malsburg-type synapses this mechanism must be fast enough to allow rapid changes.

The NMDA Receptor-Ionophore Complex

One (but probably not the only one) such mechanism that could fulfill both conditions has been described. Much recent data indicate that the NMDA receptor channel complex may serve as the molecular mechanism for Hebb-type synapses. This channel is voltage-dependent (for a recent review see Cotman & Monaghan, 1988), i.e. it opens under two conditions: (a) presynaptic activity that releases a transmitter that binds to the receptor, and (b) sufficient postsynaptic depolarization. At membrane potentials around resting potential, NMDA channels are blocked by magnesium ions; if the membrane is depolarized, to say -30 to -20 mV, the block is removed and the channel is switched on. The NMDA receptor-linked channel is permeable to Na^+, K^+ and Ca^{2+}. Thus, NMDA receptors provide a mechanism to (a) amplify synaptic currents, and (b) increase Ca^{2+} influx *if* a particular level of postsynaptic depolarization is present. Ca^{2+} is hypothesized to trigger a number of Ca^{2+}-dependent molecular changes inside the postsynaptic terminal that modify the efficacy of synapses (other than those gated by NMDA). Some of the potential targets of Ca^{2+} are shown in Figure 2 (for a recent review see Kennedy, 1989). It can be assumed, however, that some of these processes are rapid enough to allow fast changes in weight; in particular, if one assumes that the learning unit is the dendritic spine rather than the whole cell. Furthermore, the presence of the NMDA receptor-ionophore complex in neural nets would also enable the formation of assemblies *without* permanent changes in synaptic weight. As mentioned above, the NMDA receptor amplifies synaptic currents and, hence, the firing properties of the postsynaptic neuron, if it detects a correlation between signals converging on one cell. In a population of neurons cross-connected by such correlation-detecting, NMDA-mediated connections, those neurons that are active at or near the threshold level (e.g. they are simultaneously activated by different components of a stimulus) will form positive feedback loops and persist as a co-operative, coherently active subpopulation. Thus the formation of

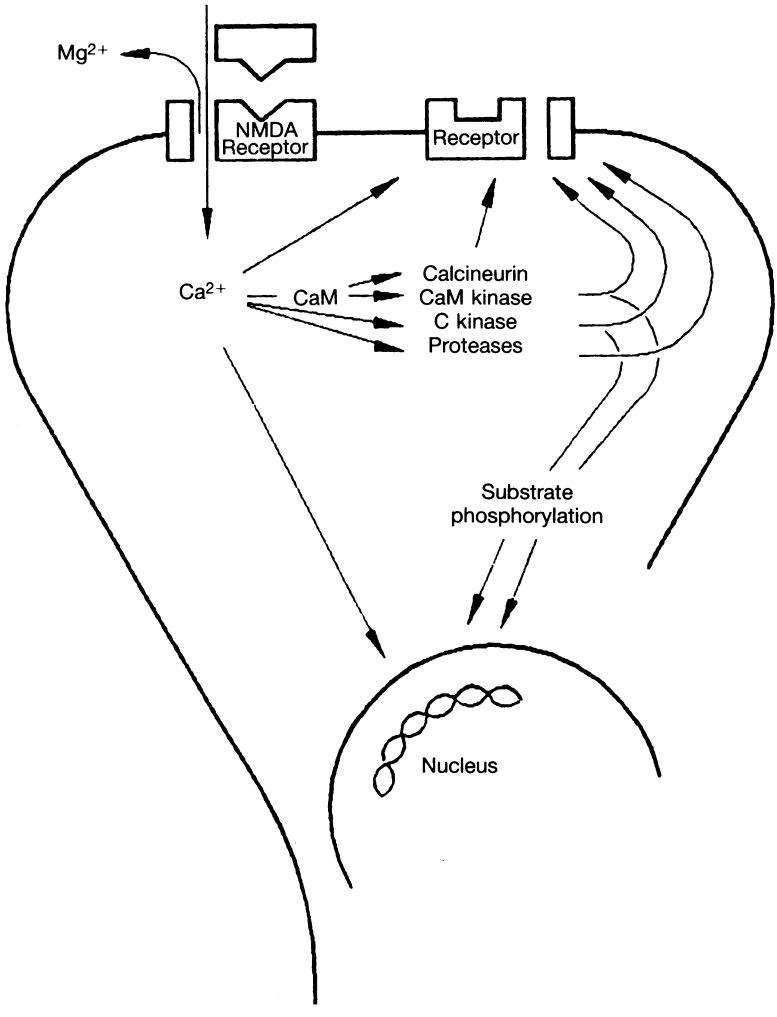

FIGURE 2. Schematic diagram of possible targets of Ca^{2+}. Ca^{2+} regulates the function of channel proteins and several cytosolic proteins including protein kinases, phosphatases (calcineurin) and proteases which, in turn, affect the excitability of the postsynaptic membrane. It also influences directly and indirectly neuronal protein synthesis.

(transient) assemblies is not dependent on changes in synaptic efficacy, as initially supposed by Hebb and von der Malsburg. It appears that the NMDA receptor can control the production of different types of assemblies.

Awareness Is a Function of the Capacity to Represent

Central to the present hypothesis is that synaptic modifications depend on the depolarization of the postsynaptic membrane reaching a certain threshold. The formation of assemblies in a network is therefore dependent on the degree of activation of its neurons. The answer to the question posed above concerning the specific conditions for the occurrence of consciousness brought about by the convergence of specific and unspecific afferents would therefore be as follows: this form of processing is the prerequisite for the formation of assemblies between neurons with NMDA connections; it determines the rate and the extent of their formation as well as their duration. A global activation of a cell population, as exerted by the mesencephalic reticular formation on the cortex, would generally increase the probability of assemblies being formed. Focal, 'searchlight'-like changes, e.g. from the dorsal thalamus, would increase the probability of assembly formation only in one cortical domain. 'Patterned' searchlights would relatively increase the probability of certain classes of assemblies being formed. Conscious and unconscious brain structures therefore differ in the rate at which assemblies can be generated. Common to all forms of unconsciousness—focal or general—is a lowered production rate. Attention is characterized by an increased rate. The rate at which the different types of assemblies are generated determines:

1. the amount of representations generated in a given period of time;
2. the maximal size of assemblies that can be generated in a given time period and, hence, the maximal complexity of representations. The degree of activation of a neural net therefore determines *qualitative* differences in what the brain can and cannot represent;
3. the duration of representations;
4. the amount, complexity and duration of metarepresentations referring to internal states that can be formed per time; and
5. the amount and the quality of computational processes possible in a given period of time—for these depend on the amount and the quality of given representations.

My hypothesis is that deficits in phenomenal consciousness as well as certain functional deficits occur if the rate of assembly formation falls below a critical threshold level. Whenever this level is surpassed, phenomenal states must necessarily occur. A critical production rate of neural assemblies is the necessary and sufficient condition for the existence of phenomenal states. Awareness is the result of the system's capacity to actively generate representations and metarepresentations.

Experimental Tests

This hypothesis is testable, and I briefly describe two groups of experimental findings that, in my opinion, support it.

1. One prediction is that focused attention increases the formation rate of assemblies. If focused arousal in a specific sensory or multisensory circuitry is induced in human or animal experiments it is accompanied by characteristic event-related EEG changes (for a recent review see Sheer, 1989). Typically, the arousing signal gives rise to a so-called 40 Hz EEG, this designation being a shorthand notation for different narrow-frequency bands in the high-frequency gamma range (35–85 Hz). These large-amplitude, highly synchronous bursts have been observed in different focused arousal paradigms, different species and different sensory modalities, such as the olfactory, visual and auditory systems. The 40 Hz EEG is significantly reduced in amplitude under anesthesia (Galambos, Makeig, & Talamachoff, 1981). The explanation for these EEG changes is that large groups of neurons in different areas have begun to fire coherently; an assembly has been formed as a result of the interaction between specific sensory afferents and the phasic arousal system which has its origin in the parabrachial nucleus of the pons. The neural substrate of focused attention itself is the coherent activity of assemblies in specific or multisensory circuitry.

2. NMDA-gated synapses do not appear to be involved in signal propagation in networks with rigid synapses; rather, their function seems to be limited to the control of activity-dependent plastic processes. Selective blockade of this mechanism should therefore prevent only plastic changes but not other processes involving 'conventional' synaptic transmission. According to the present hypothesis, this intervention should lead to disturbances of phenomenal consciousness. This is in fact the case. A number of selective antagonists of NMDA receptors are known. These include the anesthetics phencyclidine and ketamine. Both compounds produce general anesthesia. Their profile as anesthetics differs markedly from other conventional general anesthetics, such as barbiturates, ether, chloroform or halothane. Their properties are characterized by profound analgesia, maintenance of most protective reflexes, poor muscle relaxation and lack of cardiorespiratory depression. They do not produce a true sensory blockade at spinal or brain stem levels. Studies on visual and somesthetic evoked potentials indicate that sensory input reaches the primary sensory cortex, but that the subsequent processing in association areas is depressed (Chen, Ensor, Russel, & Bohmer, 1959; Luby, Cohen, Rosenbaum, Goltlieb, & Kelley, 1959). In subesthetic doses NMDA antagonists cause altered states of consciousness (Pollard, Uhr, & Stern, 1965). Patients anesthetized with ketamine or phencyclidine or phencyclidine addicts report that they experience vivid dreams, sensory illusions, visual and auditory hallucinations, distortion of the body image

and disorganized thought. Phencyclidine may produce psychotic episodes closely resembling certain forms of schizophrenia. Many other NMDA antagonists like cyclazocine, pentazocine, dextrorphan, dextromethorphan induce similar perceptual abnormalities.

What Can This Hypothesis Explain?

In my opinion, this hypothesis is of explanatory value. It offers an explanation for a number of physiological, psychological and philosophical aspects of consciousness, some of which are introduced briefly in the following list.

1. The hypothesis can explain the existence of different forms of disturbances of consciousness as a disturbance of one particular category of brain processes. Lowered probability of assembly formation is the common denominator of widely differing disturbances of consciousness. Global disturbances occur if the rate of formation of neural assemblies generally falls below a critical value. Focal disturbances, on the other hand, occur if this threshold is not reached, either in certain areas of the brain or for certain configurations of assemblies. A lowered probability of assembly formation also constitutes the common link shared by all the diverse causal pathways leading to states of unconsciousness (e.g. by various brain lesions, pharmacological intervention, various forms of anesthesia, oxygen lack, etc.).

2. It can explain the phenomenal aspect of attention. Objects to which attention is directed, not only elicit behavioral reactions, but are also different subjectively. According to the present interpretation, this occurs in those domains in which the assembly turnover is enhanced by thalamocortical activation.

3. It can explain *quantitative* aspects of consciousness. Extent and intensity of phenomenal states above the critical threshold are determined by the amount of representations generated per time.

4. It can explain *qualitative* changes of consciousness. Pharmacologically induced states of altered consciousness are often caused by transmitter-like substances which interact with certain central synapses. If such pharmacological intervention blocks or favors the participation of certain groups of neurons in the formation of an input-dependent assembly, identical input patterns would lead to different assemblies. The same sensory stimulus configuration would give rise to 'abnormal' representations. The same applies to assemblies generated by internal stimulus configurations. The hypothesis would thus explain illusions and hallucinations.

5. It can explain the close connection between the existence of certain phenomenal states and certain functional states. Phenomenal states and functional states dependent on consciousness would be that fraction of all brain states, the occurrence of which is dependent on large amounts and/or

great complexity of representations. Qualia and functional states are based on the same instantiation. In this view, phenomenal consciousness would not be an epiphenomenon. It also follows from the present hypothesis that the intentional contents of mental states likewise do not represent an epiphenomenon, but are real, causally effective states. Central to Hebb's ideas is that plastic neural nets can function as coincidence detectors, and that the coactivation of certain neuron populations by stimulus patterns creates a representation of those patterns. Hence, coincidence detection structures representations. By these means the relational features in given stimulus patterns (of external or internal origin) become structural causes in Dretske's sense. In my opinion, Dretske has provided no answer to the question of how intentional contents can be the cause of physical processes. Instead, he has circumvented the question by relegating it to the history of information processing systems, their evolution or their generation by a learning process. He does not tackle the question of how intentionality can be created in the genesis of such systems. The beginning of a solution offered by the present hypothesis (for learning systems) is that coincidence detection is the general principle by which content is implemented and becomes a structural cause in Dretske's sense. In the present context, I do not elaborate on this important question. It is closely related to the problem under discussion but not decisive for the exposition of the present hypothesis.

6. It can explain the relationship between consciousness and certain forms of memory. On the basis of current physiological knowledge, it can be assumed that a continuity exists between transient and permanent assemblies. Ca^{2+} initiates both short-term and long-term synaptic changes. It is conceivable that the life span of an assembly is dependent on the amount of second messenger produced. This could explain the well-known fact that in agnosias loss of conscious perception of a particular stimulus category is coupled with a loss of memory for the same category (Saper & Plum, 1985). Ketamine and phencyclidine both induce severe retrograde amnesia for events perceived before and during intoxication. By contrast, the specific drug-induced hallucinations are remembered.

7. It explains the occurrence of subjectivity within a system as the necessary result of representational and metarepresentational processes. Nerve nets with a high rate of assembly formation can produce more and more complex, and thus qualitatively different, representations than nets with a lower formation rate. Nets with a high formation rate will automatically generate active metarepresentations of internal states, the complexity of which will be limited by the complexity of the physical tokens generated per time. At a sufficiently high formation rate such systems will develop self-referential, introspective, metacognitive activities. In such systems an inner perspective will automatically develop. Subjectivity arises necessarily in nerve nets with high rates of assembly formation.

8. It may explain the ineffability of phenomenal states. Paradoxically, if introspective consciousness is directed to any other state of consciousness it extinguishes and transcends that state of consciousness. As Dennett (1988) has pointed out, this is the reason for the unavoidable ambiguity in attempts to say what qualia really are. The proposed representational system has the same property: assemblies cannot represent themselves in the act of representing. If neural systems 'turn to themselves', new assemblies, different in form and shape, will replace the initial representations. According to this view, ineffability has a *physical* cause.

What Is It Like to Know What It Is Like to Be a Bat?

Have we now escaped from the clutches of the classical qualia-are-irreducible-dualistic-aspects-of-the-mind argument? It was not the primary intention of this paper to discuss the familiar arguments once again, rather, to present an empirically testable thesis that might narrow the gap: physiology can explain what are the physically concrete conditions in which subjective states occur. Evidently, the present hypothesis provokes both, neurophysiological and philosophical, questions. Its impact on the different arguments against a physicalistic account of the phenomenal aspects of mentality remains to be investigated in detail. In my opinion, it will lead to the disappearance of the apparent dilemma. If subjectivity emerges from a system's representational and metarepresentational activity, it will be possible to describe these activities—the formation of representations and their interactions—in a physical framework. It will be possible to translate events in alien representational systems into activities of our own representational system. We would learn to manipulate tokens that represent representations in other representational systems. In doing so, we would not exactly feel like bats (or another alien system) ourselves, but we would be able to have a theory of what it is like to be a bat, i.e. we would *know* it.

With a theory of subjective states as manifestations of the representational activity of a system it turns out that Nagel's problem has a natural solution. It is a question of (meta)representing active representational systems in other representational systems. That this procedure does not lead to identical subjective states in both the observed and the observing representational system is trivial.

Notes

1. An earlier version of this article was presented at the conference on 'The Phenomenal Mind: How Is It Possible? Why Is It Necessary?', Center for Interdisciplinary Studies, University of Bielefeld, Bielefeld, Germany, 16 May 1990.

2. I share van Gulick's opinion that the word 'phenomenal' should not be used to refer exclusively to sensory qualia or raw feels: 'it is a serious mistake to equate the phenomenal aspect of mind solely with such properties. We should not forget that the idea of the phenomenal structure of experience entered philosophical thought through Kant who introduced it in the context of rejecting the sensational theory of experience associated with traditional empiricism. Phenomenal experience is not merely a succession of qualitatively distinguished sensory ideas, but rather the organized cognitive experience of a world of objects and of ourselves as subjects within that world. To focus exclusively on raw feels would be a mistake in at least two respects: 1. it would provide too narrow a definition of what needs to be explained, and 2. I doubt that qualia and raw feels can themselves be understood in isolation from the roles they play within the richer Kantian structure of phenomenal experience' (van Gulick, 1990).

References

Bienenstock, E. (1983). Cooperation and competition in central nervous system development: A unifying approach. In E. Basar, H. Flohr, H. Haken, & A.J. Mandell (Eds.), *Synergetics of the brain* (pp. 250–263). Berlin: Springer-Verlag.

Bieri, P. (1987). Pain: A case study for the mind–body problem. *Acta Neurochirurgica*, *38*, 157–164.

Chen, G., Ensor, C.R., Russel, D., & Bohmer, G. (1959). The pharmacology of 1-(1-phencyclohexyl piperidine HC1). *Journal of Pharmacology and Experimental Therapeutics*, *127*, 241–250.

Churchland, P.M. (1985). Reduction, qualia and the direct inspection of brain states. *Journal of Philosophy*, *82*, 8–28.

Cotman, C.W., Monaghan, D.T. (1988). Excitatory amino acid neurotransmission: NMDA receptors and Hebb type synaptic plasticity. *Annual Review of Neuroscience*, *11*, 61–80.

Crick, F. (1982). Do dendritic spines twitch? *Trends in Neurosciences*, *5*, 44–46.

Dennett, D.C. (1988). Quining qualia. In A.J. Marcel, & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 42–77). Oxford: Clarendon Press.

Farrell, B.A. (1962). Experience. In V.C. Chappell (Ed.), *The physiology of mind* (pp. 23–48). Englewood Cliffs, NJ: Prentice Hall.

Feigl, H. (1958). The 'mental' and the 'physical'. In H. Feigl, M. Scriven, & G. Maxwell (Eds.), *Concepts, theories and the mind–body problem* (Minnesota Studies in the Philosophy of Science, Vol. 2, pp. 320–492). Minneapolis: University of Minnesota Press.

Flohr, H. (1989). Schwierigkeiten der Autocerebroscopie. In E. Pöppel (Ed.), *Gehirn und Bewußtsein* (pp. 61–71). Weinheim: VCH Verlagsgesellschaft.

Galambos, R., Makeig, S., & Talamachoff, P. (1981). A 40 Hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Science (USA)*, *78*, 2643–2647.

Hebb, D.O. (1949). *The organization of behavior*. New York: Wiley.

Heilman, K.M., Valenstein, E., & Watson, R.T. (1985). The neglect syndrome. In

J.A.M. Fredericks (Ed.), *Handbook of clinical neurology: Vol. I(45): Clinical neuropsychology* (pp. 153–183). Amsterdam: Elsevier.

Jackson, F. (1982). Epiphenomenal qualia. *American Philosophical Quarterly*, *32*, 127–136.

Kennedy, M.B. (1989). Regulation of neuronal function by calcium. *Trends in Neurosciences*, *12*, 417–420.

Kohonen, T. (1983). Representation of information in spatial maps which are produced by self-organization. In E. Basar, H. Flohr, H. Haken, & A.J. Mandell (Eds.), *Synergetics of the brain* (pp. 264–273). Berlin: Springer-Verlag.

Kohonen, T. (1984). *Self-organization and associative memory: Springer series in information sciences, Vol. 8*. Berlin: Springer-Verlag.

Lewis, D. (1983). Extrinsic properties. *Philosophical Studies*, *44*, 197–200.

Loar, B. (1990). Phenomenal properties, *Philosophical Perspectives*, *4*, 81–108.

Luby, E.D., Cohen, B.D., Rosenbaum, G., Gottlieb, J.S., & Kelley, R. (1959). Study of a new schizophrenomimetic drug. *Archives of Neurology and Psychiatry*, *81*, 363–364.

Lycan, W.G. (1987). *Consciousness*. Cambridge, MA: MIT Press.

Lycan, W.G. (1990). What is the subjectivity of the mental? *Philosophical Perspectives*, *4*, 109–130.

Meehl, P.E. (1966). The compleat autocerebroscopist: A thought experiment on Professor Feigl's mind–body identity thesis. In P.K. Feyerabend & G. Maxwell (Eds.), *Mind, matter and method* (pp. 103–180). Minneapolis: University of Minnesota Press.

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 435–450.

Nagel, T. (1986). *The view from nowhere*. Oxford: Oxford University Press.

Pollard, J.C., Uhr, L., & Stern, E. (1965). *Drugs and phantasy: The effects of LSD, psilocybin and sernyl on college students*. Boston: Little Brown.

Ritter, H., Kohonen, T. (1989). Self-organizing semantic maps. *Biological Cybernetics*, *61*, 241–254.

Saper, C.S., & Plum, F. (1985). Disorders of consciousness. In J.A.M. Fredericks (Ed.), *Handbook of clinical neurology, Vol. I (45): Clinical neuropsychology* (pp. 107–128). Amsterdam: Elsevier.

Sheer, D.E. (1989), Sensory and cognitive 40-Hz event-related potentials: Behavioral correlates, brain function, and clinical application, In E. Basar & T.H. Bullock (Eds.), *Brain dynamics. Progress and perspectives: Springer series of brain dynamics, Vol. 2* (pp. 339–374). Berlin: Springer-Verlag.

van Gulick, R. (1990, May). *Understanding the Phenomenal Mind: Are We All Just Armadillos*? Paper presented at the conference on 'The Phenomenal Mind: How Is It Possible? Why Is It Necessary?' Center for Interdisciplinary Studies, University of Bielefeld, Bielefeld, Germany.

von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, *14*, 85–100.

von der Malsburg, C. (1981). The correlation theory of brain function. *Internal Report*, 81–2. Department of Neurobiology, Max Planck Institute for Biophysical Chemistry, Göttingen.

von der Malsburg, C. (1983). How are nervous structures organized? In E. Basar,

H. Flohr, H. Haken, & A.J. Mandell (Eds.), *Synergetics of the brain* (pp. 238–249). Berlin: Springer-Verlag.

Wilshaw, D.J., von der Malsburg, C. (1976). How patterned neural connections can be set up by self-organization, *Proceedings of the Royal Society (London), Series B, 194*, 431–445.

HANS FLOHR is professor of neurobiology at the University of Bremen and director of the Brain Research Institute, Bremen, Germany. His research is centered on (1) the cellular and molecular mechanisms of learning and memory and (2) the physiological correlates and conditions of consciousness. He is the editor of *Post-lesion Neural Plasticity* (Springer Verlag, Berlin/Heidelberg, 1988).