April 3, 1996 LBL-37944new

Chance, Choice, and Consciousness: A Causal Quantum Theory of the Mind/Brain *

Henry P. Stapp
Lawrence Berkeley Laboratory
University of California
Berkeley, California 94720

Abstract

Quantum mechanics unites epistemology and ontology: it brings human knowledge explicitly into physical theory, and ties this knowledge into brain dynamics in a causally efficacious way. This development in science provides the basis for a natural resolution of the dualist-functionalist controversy, which arises within the classical approach to the mind-brain system from the fact that the phenomenal aspects are not derivable from the principles of classical mechanics. A conceptually simple causal quantum mechanical theory of the mind/brain is described, and used to examine the necessity and function of consciousness in brain process.

Invited Paper for the Conference:
Toward a Scientific Theory of Consciousness 1996
University of Arizona, Tucson, April 8-13, 1996.

---

*This work was supported by the Director, Office of Energy Research, Office of High Energy and Nuclear Physics, Division of High Energy Physics of the U.S. Department of Energy under Contract DE-AC03-76SF00098.

Disclaimer

1. Introduction: Knowledge in the Quantum Formalism.

Classical Mechanics purports to describe the physical world and how it functions, and claims to achieve this goal without bringing in thoughts, feelings, or any other experiential aspect of nature. For centuries this restriction to the purely physical was regarded as an important virtue of science: science had taken us beyond the primitive superstition that spirits were lurking everywhere, and causing things to happen. Instead, the physical world was asserted to be built out of nothing but quantifiable properties that could be localized in a spacetime, and whose functioning was completely determined by rigid mathematical laws that referred to nothing but these physical properties themselves. Thus when the creators of quantum theory introduced human experience and "our knowledge" into the theory of atomic phenomena their move was initially opposed by the scientific community. Soon, however, this important enlargement of the scientific conception of basic physical theory came to be accepted, at least nominally, by most workers in the field.

Recently some quantum theorists have been trying to exorcise "the observer" from quantum theory. These attempts encounter difficulties, which I shall mention later. But in any case the important point in our quest for a scientific theory of consciousness is not that the basic physical theory *might* some day be able to be formulated without introducing observers. It is rather that our basic theory of matter, in its contemporary orthodox form, has an explicit and dynamically efficacious place for conscious experiencings.

This focus of orthodox quantum theory on the experiential aspects of nature was emphasized in the opening words of Niels Bohr's principal book on the subject, Atomic Theory and the Description of Nature:

"The task of science is both to extend the range of our experience and reduce it to order."[1]

Later he says:

"In physics... our problem consists in the co-ordination of our experience of the external world ... "[2]

and

"In our description of nature the purpose is not to disclose the real essence of phenomena but only to track down as far as possible relations between the multifold aspects of our experiences."[3]

An analogous statement by Heisenberg is:

"The conception of the objective reality of the elementary particles has evaporated in a curious way, not into the fog of some new, obscure reality concept, but into the transparent clarity of a mathematics that represents no longer the behaviour of the elementary particles but rather our knowledge of this behavior."[4]

As these quotations indicate, the original formulation of quantum theory was essentially epistemological rather than ontological: it was about "our knowledge" of the physical world, rather than the physical underpinnings of this knowledge. This epistemological formulation was subsequently "ontologicalized" by von Neumann and Wigner, who used, in effect, Heisenberg's notion (borrowed from Aristotle) of 'potentia', which means 'objective tendencies' for things to happen. Another term is 'propensities'. In this ontological form the physical and experiential aspects of nature are bound together in a tight dynamical structure. Our knowledge itself becomes explicity represented in the theory, and is tied into the physical substrate in a dynamically efficacious way: Epistemology and ontology become united.

This dynamical joining of knowing and being is just what a theory of the mind/brain ideally ought to achieve: it should bring phenomenal knowledge explicitly into the physical theory as an efficacious dynamical element.

Yet how can a physical theory encompass something so different from matter as our experiencings of the world.

To see how this works let us first recall how it is done in classical mechanics. All statements in science must be transcribed into statements about "our possible experiences" before they can be tested by human beings, or used make predictions about what our future experiences will be. All such prediction are based upon some prior knowledge of the world about us. But this prior knowledge never determines the state of the world exactly. In classical statistical mechanics this prior knowledge, call it K, is represented by a 'probability density function', $D(x, p; K)$. Here the argument x represents the positions of all of the particles of the system being examined, and p represents the momenta (or velocities) of these particles. The function $D(x, p; K)$ defines the probability density in phase space (i.e., in (x, p)-space) corresponding to the prior knowledge K, For example, one might know only that some set of particles of interest lie in a certain box, and have a certain temperature. This knowledge K can be represented by a particular probability density function $D(x, p; K)$

Suppose I now ask the following question: Given the statistical information represented by the probability function $D(x, p; K)$, what is the probability $P(m, e; K)$ that if I observe the system in manner m I will have experience e.

Let $E(m, e; x, p)$ be the probability (density) that if the system is in the state specified by the point (x, p) in phase space, and I observe it in manner m, then I will have experience e. Folding together these two probabilities one obtains the basic formula

$$P(m, e; K) = \int dxdpE(m, e; x, p)D(x, p; K).$$

This same formula holds in quantum mechanics. But in quantum mechanics the functions $D(x, p; K)$ and $E(m, e; x, p)$ are complex numbers, rather that positive numbers, and a classical probability interpretation is ruled out. Also, the equation motion in quantum theory is such that the different members of what in classical mechanics would be a 'statistical ensemble' of independently moving points (x, p) in phase space do not evolve independently: in quantum theory these 'independent components' are influenced by their neighbors. Thus they cannot be conceived of as members of a classical statistical ensemble. However, the function $D(x, p; K)$, taken as a whole, can be interpreted as representing, under condition K, the 'propensity' for experience e to occur if an observation in manner m is performed.

In the original Copenhagen interpretation of Bohr and Heisenberg, and their colleagues, the variables x and p referred to an external system that was being examined by the scientists. Thus in the function $E(m, e; x, p)$ the pair of variables (x, p) referred to one system (some small 'observed' part of the universe) and the pair of variables (m, e) referred to things associated with a different part of the universe, namely the brain of the observer plus his body, extended to include his measuring devices.

This mapping function connects two different parts of the universe, and therefore depends upon a separation of the physical universe into parts. But this separation of the physical world into parts was not well defined within the theory. Thus the theory, as originally formulated, was not interpretable as a basic theory of nature herself: it was fundamentally epistemological rather than ontological, as its creators repeatedly emphasized.

John von Neumann[5] and Eugene Wigner[6] ontologicalized the original Copenhagen form of the theory by identifying 'the system' with the entire universe. Then the variables x and p become the variables needed to describe the entire universe, including, in particular, the brains of the observers. The function $E(m, e; x, p)$ then relates an "experiential space", whose elements are labelled by the variables m and e, to a physical space that includes, in particular, the brain of the relevant observer.

The functions $D(x, p; K)$ and $E(m, e; x, p)$ can be regarded as the matrix elements $< x|D(K)|p >$ and $< p|E(m, e)|x >$ of *operators* $D(K)$ and $E(m, e)$. Then a huge difference between quantum mechanics and classical mechanics is that the occurrence of the experience (m, e) is identified as an actual event that acts back on the physical world: The world before this event, which is represented by $D(K)$, is transformed as follows:

$$D(K) \Longrightarrow D(K; m, e) = CE(m, e)D(K)E(m, e),$$

where C is a normalization constant. This transformation represents a 'reduction of the state': the state $D(K)$ prior to the actual experiential event that is represented by $E(m, e)$ is transformed to a new state $D(K; m, e)$ that incorporates the new conditions labeled by (m,e). In keeping with this meaning, the operator $E(m, e)$ satisfies the (idempot) condition

$$E(m, e)E(m, e) = E(m, e):$$

a single experience acting twice has the same effect as its acting once.

This description shows how our experiencings become woven into the fabric of the quantum mechanical description of nature: they are the identifiers of the actual events that are the coming into being of these experiencings, and that act efficaciously upon the mathematical structure that represents the physical aspect of nature. This physical aspect constitutes the more subtle aspect of reality: it is a substrate of *tendencies* for actual events to occur.

Human experiences are presumed to be very high-level forms of actual events: we focus here on these special kinds of events because our topic here is consciousness, because they are the foundation of our science, and because we have direct information about them.

Quantum mechanics is predicated on the fact that our experiences of the physical world—our immediate phenomenal knowledge of it—can be described in the language of classical mechanics, considered as an extension of ordinary every-day language. Quantum dynamics itself, in the von Neumann/Wigner form, entails that in most situations classical mechanics gives a description of the phenomena that is difficult to distinguish, empirically, from the one given by quantum mechanics. In particular, quantum mechanics entails that it will be only after we have delved fairly deeply into the dynamical details brain processing that we will be able to discern empirically the difference between the quantum description of the brain dynamics and certain essentially classical ones, even though the ontological differences are enormous.

2. Inadequacy of the Classical-Mechanical Description of the Brain.

The classical-mechanical description of the physical world (although empirically false) is, logically speaking, dynamically complete, even though it never mentions the experiential (i.e., phenomenal) aspects of nature. To account, within this framework, for the factual occurrence of these experiential aspects some scientists and philosophers have been led to suppose that certain brain activities simply 'evoke' corresponding experiences without the latter reacting back on the brain. According to this idea, the experiential world is merely an epiphenomenal add-on to a physical world that is completely described by classical mechanics.

This scenario might be logically possible, but it seems preposterous that nature should create a whole extra world, which is totally unlike the physical world, and in no way entailed by the laws that govern the physical world, and then give this add-on world no dynamical role to play.

The unnaturalness and non-parsimoniousness of this (classical dualistic) notion has led to an opposing (classical identity/functionalist) claim that experiencings simply ARE certain functional activities of the brain , described in a phenomenal rather that physical/functional language: i.e., that all that there is in nature (at a level that is adequate to cope with the mind-brain problem) are the classically described physical/functional activities, but that certain of these activities have also a second kind of description: a phenomenal description.

To examine this claim, suppose brain science has finally evolved to the point where it can give a complete description of brain process: suppose it can provide a detailed understanding of how 'memory tracks' are laid down in the brain, and how these memory tracks are accessed by later brain activities. And suppose that the brain scientist could even wire up a brain and map out its various patterns of activity in sufficient detail to be able to follow through what happens in the brain when the person is asked "What did you eat for breakast this morning?" Suppose the brain scientist is able to follow through the progression of patterns of excitations and see how the physical memory tracks laid down during breakfast come to be accessed, and to see how the content of those memory tracks feeds into the process that finally produces the spoken reply "Ham and Eggs!", and even the reply to the follow-up question "What color were the eggs?": "They were a chalky-whiteish kind of yellow, rather than an orange-ish shade of yellow!" And suppose the brain scientist's description is detailed enough to see even the laying down of certain memory tracks that will allow the person to respond to later queries about his sequence of thoughts as he was formulating his answers to the questions. Suppose further that the brain scientist is able to construct a mapping from the physical space of certain kinds of patterns of neural activity to corresponding 'phenomenal events' described in a phenomenal language, and that this mapping is such that it fits perfectly with all the responses that that person makes to questions about his "experiencings" of pain, of color, and of every other kind of experience that he says he has.

Within this context we consider, in the framework of a classical-mechanical conceptualization of the brain, the two alternative claims:

1. The phenomenal activity IS EVOKED BY a neural/functional activity.
2. The phenomenal activity IS the corresponding neural/functional activity.

The advocate of claim 2, which I call (classical) functionalism, can claim parsimony, and can point to the unnaturalness of the existence, asserted by claim 1, of a whole world that is fundamentally different from the physical world, and has no effect upon the physical world. The existence of such a world would seem to require a whole new machinery in nature, a machinery that would somehow 'cause' the phenomenal events to occur, or 'evoke' them, even though nothing in the classical physical laws requires the existence of any such extra machinery.

On the other hand, the advocate of claim 1, which I call (classical) dualism, can insist that any claim that two different descriptions describe one single thing must be supported by some explanation of how the one thing acquires these two descriptions. For example, the claim that temperature in thermodynamics IS the same as mean kinetic energy in statistical mechanics is supported by the fact one can *deduce* this correspondence from the laws of physics. The claim that The Morning Star IS The Planet Venus, is also explained on the basis of the laws of physics, by noting that the phrase "The Morning Star" has an original meaning that refers to a certain kind of 'experiencing' (an experiencing of a certain brightness in the morning sky not too far from the horizon), and by the fact that this experiencing can be deduced, on the basis of the laws of physics, to be caused by sunlight reflected off of the planet called Venus, provided such 'experiencings' of an observer can be assumed to be evoked by corresponding activities in the brain of this observer. But in the case of the claim that "The Pain P IS The Functional Brain Activity F" there is no possibility of deducing this connection from the orthodox principles of classical mechanics.

The functionalist can reply that he has in fact, on the basis of the principles of classical mechanics, provided a detailed causal account of the very activity in the brain that constitutes experiencing. He can claim that his phenomenal knowings ARE, precisely, his brain's knowings of certain aspects of itself. He can claim that his experiencings ARE his brain's functional activities of laying down and retrieving certain kinds of memory tracks that: 1) contain all of the information that he feels that he is becoming aware of; and 2) initiate all of the things that he feels he is initiating.

The dualist can reply that classical mechanics allows all physically describable behavior to be deduced in principle, but does not itself imply that any ingression or flow of information, or any other physical or functional activity, IS a felt experiencing of the kind that we are immediately acquainted with. Thus the functionalist claim constitutes *adding*, to what classical mechanics itself entails, a certain kind of reality that classical mechanics does not strictly entail, namely our felt experiencings. Logically, within the framework defined by classical mechanics plus add-ons, the phenomenal feelings are add-ons, added on to account for realities that we know exist. But these realities cannot be *identical* to anything entailed by classical mechanics because classical mechanics does not entail them. In brief, if CM implies the existence of F but not the existence of an immediately known reality P, then within a logical framework that includes CM the reality P cannot be claimed to be identical to F.

The functionalist claim is therefore not logically compatible with the acceptance of classical mechanics as an adequate framework for describing brain dynamics. On the other hand, the classical dualist position entails the causal inertness of our thoughts and feelings.

The dilemma arises precisely because there is something that is known first-hand to be ontologically real and present, but whose presence is not implied by the physical theory that is being used to describe the system. This egregious omission arises from a faulty conceptualization of the situation, namely from the use for the conceptualization of the mind/brain system of a physical theory that has been found by science to be inadequate precisely at the point at issue, namely the relationship between our physical description of nature and our phenomenal knowledge of it.

Within the von Neuman/Wigner quantum ontology the phenomenal facts are introduced from the beginning as basic actualities, and they are, in fact, the very actualities that orthodox quantum theory is about. And these actualities are causally efficacious. Hence within this quantum description of nature there is no need to introduce into the theory any element not clearly entailed by the basic principles, or any element that is dynamically inert: we are not forced by science accept an unreasonable ontological stance on either point, provided we accept what basic science has been telling us for seventy years.

3. Macroscopic Quantum Effects In Brain Dynamics.

Brain Dynamics is controlled by chemical processes. Eventually, it will probably become important to have a coherent account in which these chemical processes fit seamlessly into the whole process. Hence ultimately, for these basically chemical reasons, a quantum description will presumably be needed

But I wish to focus here on more macroscopic quantum effects: effects that would distinguish the quantum treatment from a classical model in which the currents flowing along neurons are described in classical terms.

Brain process is essentially a search process: the brain, conditioned by earlier experience, searches for a satisfactory response to the new situation that the organism faces. It is reasonable to suppose that a satisfactory response will be progammed by a template for action that will implemented by a carefully tuned pattern of firings of some collection of neurons. This executive pattern would be a quasi-stable vibration that would commandeer certain energy resources, and then dissipate its energy into the initiation of the action that it represents. [cf. Ref. 7 for a more detailed description of this process.]

If the programmed action is complex, and refined, then this executive pattern must contain a great deal of information, and must, accordingly, be confined to a small region of phase space. Stated differently, the relative timings of the pulses moving along the various neurons, or groups of neurons, will have to conform to certain ideals to within very fine levels of tolerance. The problem is: How does the hot wet brain, which is being buffeted around by all sorts of thermal and chaotic disturbances, find its way to such a tiny region in a timely manner?

The problem, reduced to it basic dynamical form, is this: How does a point in a 3n-dimensional space (where n represents some huge number of degrees of freedom of the brain) which is moving in a potential well that essentially blocks out those brain states that are not good solutions to the problem (i.e., that do not represent templates for satisfactory actions, under the conditions at hand), that but does not block the good solutions, find a good solution, under chaotic initial conditions. Classically, the point in the 3n-dimensional space must just follow the deterministic equation of motion until it eventually wiggles its way out of the potential well. But the quantum system has some effective chaos built in, and it effectively explores all possible ways to get out, simultaneously. Moreover, this cloud it virtual possibilities satisfies an essentially hydrodynamic equation of motion: it acts like a single glob of water, rather than like a collection of independently moving droplets. [See Feynman, Ref. 8]. That is, the motion of each point in the cloud is influenced by its neighbors, as was emphasized earlier. But then when some parts of the glob find their way out of the potential well, and thus flow out of the confined region leaving a partial void, the nearby parts of the glob will flow in to take their place, and will then in turn flow out. Thus all of the glob will quickly flow out, like water flowing out of a leaky bucket.

The brain is a quantum system, and will automatically use this hydrodynamical property, and hence will undoubtedly operate faster in searching for an acceptable template for action than its classical counterpart can. Thus the need to use quantum theory is not just a philosophical matter: it will be needed to account for the speed of (analog) search processes.

4. Decoherence.

It has often been observed that the coupling of a system to its environment has a tendency to make interference phenomena that are present in principle within quantum systems difficult to observe in practice. Phase relationships, which are essential to interference phenomena, get diffused into the environment, and are difficult to retrieve. The net effect of this is to make a large part of the observable phenomena in a quantum universe similar to what would be observed in a world in which certain collective (i.e., macroscopic) variables are governed by classical mechanics. This greatly diminishes the realm of phenomena that require for their understanding the explicit use of quantum theory.

These decoherence effects will have a tendency to reduce, in a system such as the brain, the distances over which the idea of a simple single quantum system holds. This will reduce the distances over which the simple hydrodynamical considerations described above will hold. However, the following points must be considered.

a) A calcium ion entering a bouton through a microchannel of diameter x must, by Heisenberg's indeterminacy principle, have a momentum spread of $\hbar/x$, and hence a velocity spread of $(\hbar/x)/m$, and hence a spatial spread in time t, if the particle were freely moving, of $t(\hbar/x)/m$. Taking t to be 200 microseconds, the typical time for the ion to diffuse from the microchannel opening to a triggering site for the release of a vesicle of neurotransmitter, and taking x to be one nanometer, one finds the diameter of the wave function to be about 0.04 centimeters, which is huge compared to the 1/100000000 centimeter size of the calcium ion. There is, therefore, in brain dynamics a powerful counterforce to the mechanisms that tend to diminish quantum coherence effects. More generally, one cannot, without collapses, keep the wave function of a calcium ion confined to a region that is not huge compared to the size of the ion. This entails that classical ideas cannot be adequate: the brain must, if no collapes occur, evolve into an amorphous superposition of states corresponding to a continuum of different possible macroscopic behaviours.

b) The normal process that induces decoherence arises from the fact that a collision of a state represented by a broad wave function with a state represented by a narrow wave packet effectively reduces the coherence length in the first state to a distance proportional to the width of the second state. But in an aqueous medium in which all the states of the individual systems have broad packets this mechanism is no longer effective: coherence lengths can remain long.

c) Even if the coherence length were only a factor of ten times the diameter of the atom or ion involved in some process, the cross section involved would be a hundred times larger. The search processes under consideration here involves huge numbers of atoms and ions acting together, and the cross-section factors multiply. Thus even a small effect at the level of the individual atoms and ions could give, by virtue of the hydrodynamical effect, a large quantum enhancement of the efficiency of an essentially aqueous macroscopic search process.

5. Everett and Consciousness.

Einstein[9] illustrated the central logical problem in contemporary quantum theory with a simple example. It involves a radioactive source, a detector of some product of the decay, a pen that draws a line on a moving strip of paper and makes a blip when the decay is detected, and a human observer of the blip. If one uses the Schroedinger then one finds that the system evolves into a continuous superposition of states corresponding to all possible positions of the blip on the strip of paper. But when the human observer looks, he sees the blip in one well defined place. Thus the Schroedinger equation is not telling the whole story. If one wants to have an account of what is actually happening, then something else needs to be added, namely Heisenberg's 'transition from possible to actual' (or some substitute for it) that allows the many possibilities generated by the Schoedinger dynamics to be reduced to the single actually experienced reality. The von Neumann/Wigner form of quantum theory accepts the Heisenberg transitions from possible to actual as real events.

The strangeness of Heisenberg's idea [10] of transitions from 'possible' to 'actual', naturally has led scientists to explore diligently the possibility that these transitions never happen: that the Schroedinger equation never fails. This possibility was examined in some detail by Everett[11]. The consequence of that work, and of many later efforts to clarify it, is to focus attention even more strongly than ever on the problem of our consciousness experience. For if the

Schroedinger equation never fails then there is a huge disparity between the objective world, which is represented by the evolving state of the universe, and our subjective experiences of it.

The basic problem with this interpretation is that the needed psychological, i.e., experiential, properties of brains do not follow from the Schroedinger equation: the latter can perhaps generate independently evolving 'branches' of the wave function of the brain, with different branches corresponding to different streams of consciousness, but these branches are *conjunctively* present. However, in order to obtain the statistical predictions of quantum theory, which pertain to our experiences, the experiential streams that correspond to these different physical branches must be *disjunctive*. That is, the objective physical state will contains branch A *and* B, etc., whereas to get statistical statements about our subjective experiences one needs the logical structure of experience A *or* experience B, etc.. This means that 'mind' needs an ontology and dynamics that does not logically follow from the Schroedinger equation that controls the 'brain'. This need for a second level of reality, controlled by a dynamics that does not follow logically from the first, appears to nullify the advantage that the Everett interpretation seemed at first to provide. In any case, a spotlight becomes focused more strongly than ever on the problem of the connection between objective aspects of a mind/brain represented in the conjunctively present branches of the wave function and the disjunctive subjective aspects.

Looking at the evaluations by physicists who are pursuing environmental decoherence effects, and other essentially 'Everett' ways of approaching the problem of quantum measurement we find Zurek[12] saying, of these approaches, that they do not allow us to understand how we as 'observers' fit in, and hence they appear to him to be merely "a hint about how to proceed rather than the means to settle the issue quickly." Joos[13] says " Of course the central problem remains unsolved: Why are there local observer's?" Gell-mann and Hartle[14] emphasize that: "If history dependence can be properly introduced into the explicit treatment of quantum mechanics, then we may be able to handle individuality [of observers] with the care that it deserves". Omnes[15], who gives perhaps to most comprehensive description of these Everett-type theories says, about the Everett proposal, that he feels "it impossible to accept as a satisfactory answer to the problem of actuality." So almost forty years after the

Everett paper appeared it is acknowledged by these workers, and I think by all others who have examined the matter with sufficient care, that the problem of 'the observer' has not been banished by that approach. This issue, namely the problem of how our individual experiences fit into nature, remains the central unsolved problem, and the Everett approach makes this fact even clearer than before. The Everett approach tries to resolve the problem of the observer without going beyond the Schroedinger equation, but it certainly has not succeeded in doing so.

6. Causality and Chance.

If one adopts the quantum view that the actual events in human brains are phenomenal in character, and cause the wave functions to collapse, then one can move on to the further question of what causes a particular phenomenal event to occur. Orthodox quantum theory says 'statistical cause': the quantum state of the brain specifies the 'propensities' for the various phenomenal possibilities. Thus the cause of the phenomenal events is not the local deterministic (mechanical) sort of cause that occurs in classical mechanics. If one insists on naming what it is that picks out the one particular possibility that actually occurs in a given situation orthodox quantum theory can only answer: pure chance!

As regards the role of chance Bohr says this:

"The circumstance that, in general, one and the same experimental arrangement may yield different recordings is sometimes picturequely described as a 'choice of nature' between such possibilities. Needless to say, such a phrase implies no allusion to a personification of nature, but simply points to the impossibility of ascertaining on accustomed lines directives for the course of a closed indivisible phenomena. Here, logical approach cannot go beyond the deduction of the relative probabilities for the appearance of the individual phenomena under given conditions.[13]"

Bohr carefully avoids affirming that there actually is in nature herself an irreducible element of chance. He says, rather, that the entry of chance is due to difficulties that arise from trying to apply customary (local-reductionistic) thinking to closed indivisible phenomena. This suggests that nature herself must have some (necessarily nonlocal [See Ref. 7, p.5]) way of determining which event will actually occur: i.e., that we do not have to accept the absurdity of something definite arising out of absolutely nothing at all. But since in our causal ontology the actual events are the phenomenal events identified by the phenomenal features (m, e), these latter features must evidently enter into the selection process.

To keep the discussion from becoming too abstract at this point it is useful to introduce a simple conceivable causal quantum-mechanical model of the mind/brain.

7. General Description of Brain/Mind Dynamics.

Before going into the specific mathematical details of the model, I give a general overview. This is a brief synopsis of the description of mind/brain dynamics given in reference 7.

8.1 Body-World Schema. It is accepted here (or postulated) that there is in a person's brain a high-level represention of his body and its environment: i.e., that a person's body and its environment are represented in his brain by patterns of neurological and other brain activity. This representation in the person's brain of his body and its environment is called the 'body-world schema'. It is expanded to include representations of beliefs, and hence is sometimes called the body-world-belief schema, but I shall stick to the shorter name. Each phenomenal quantum event actualizes a body-world schema. An attentional event up-dates the body-world schema; an intentional event actualizes a body-world schema that is an image of an intended state of the body-world. It serves as a template for action.

8.2 Facilitation, Associative Recall, and Control. The persistence of a pattern of neural excitation 'etches' this pattern into the physical structure of the brain, in the sense that this pattern is 'facilitated' (made easier to activate), and that a later activation of part of the pattern tends to spread to the whole. This facilitation and spreading effect provides the basic mechanism for an explanation of associative recall, and of the control aspect of the body-world schema.

8.3 The Effect of Quantum Theory

The effect of quantum theory is essentially the same as it was in the Einstein example described earlier: the evolution controlled by the Schroedinger equation will produce, instead of one single body-world scheme, rather a continuum, consisting of a superposition of all the possibilities, with no one possibility singled out as the one that is actually experienced. Thus, for example, for every possibility in which a 'synaptic event' —the release of a vesicle of neuro-transmitter— occurs there will be other superposed possibilities in which this event does not occur; and for every situation in which an action potential spike exists at one place along an axon there will be other superposed possibilities in which the spike is a little earlier, or a little later, and still others in which it is much earlier, or much later. To extract the actually experienced reality from this amorphous conglomerate of superposed possibilities one needs, according to the Heisenberg ontology accepted here, a transition from 'possible' to 'actual'. This transition is called an actualization event: it selects and actualizes one of the alternative possibilities generated by the Schroedinger-equation-controlled evolution.

Many people, even some scientists, suppose that the quantum events occur at a microscopic level. However, there is no reason for this to be so, and no empirical evidence that is so. Indeed any evidence for microscopic quantum events would be evidence *against* the correctness or completeness of contemporary quantum theory, and no such evidence has ever been found.

The core idea of the present work is that each phenomenal event actualizes the body-world schema that specifies the image of the body-world that is experienced in that event.

8. Mathematical Formulation of the Model.

Quantum electrodynamics (extended to cover the magnetic properties of nuclei) is the theory that controls, as far as we know, the properties of the tissues and the aqueous (ionic) solutions that constitute our brains. This theory is our paradigm basic physical theory, and the one best understood by physicists. It describes accurately, as far as we know, the huge range of actual physical phenomema involving the materials encountered in daily life. It is also related to classical electrodynamics in a particularly beautiful and useful way. I take it as the basis of this work.

In this section I assume the reader to have some knowledge of the principles of quantum electrodynamics, and the notations used to describe it. I draw particularly on references [16] and [17], which describe in detail the natural connection between quantum electrodynamics and classical electrodynamics.

The brain is fairly transparent to electromagnetic radiation at frequencies that correspond to the first 10^3 standing wave modes, and the surrounding bone is a good insulator. This means that these modes should be quasi-stable. And these modes are fed by transitions between rotational states of proteins that should be activated by neuronal activity. Philip and Brian Stocklin have discussed these properties at length in connection with their own theory of consciousness.[18]

To represent the limited capacity of consciousness I assume, in this model, that the states of consciousness associated with a brain can be expressed in terms of this relatively small subset of the modes of the electromagnetic field in the brain cavity. (I also assume that events occurring outside the brain are keeping the state of the universe outside the brain cavity in a single state, so that the state of the brain can also be represented by a single state.) The brain is represented, in the method of Feynman, by a superposition of the collection of trajectories of the particles in it, with each element of the superposition accompanied by the state of the electromagnetic field that this collection of trajectories generates.

In the low-energy regime of interest here it should be sufficient to consider just the classical part of the photon interaction defined in [16]. Then the explicit expression for the unitary operator that describes the evolution from time t_1 to time t_2 of the quantum electromagnetic field in the presence of a set $L = \{L_i\}$ of specified classical charged-particle trajectories, with trajectory L_i specified by the function $x_i(t)$ and carrying charge e_i, is

$$U[L; t_2, t_1] = \exp < a^* \cdot J(L) > \exp < -J^*(L) \cdot a > \exp[-(J^*(L) \cdot J(L)/2)],$$

where, for any X and Y,

$$< X \cdot Y > \equiv \int d^4k (2\pi)^{-4} 2\pi\delta^+(k^2) X(k) \cdot Y(k),$$

$$(X \cdot Y) \equiv \int d^4k (2\pi)^{-4} i(k^2 + i\epsilon)^{-1} X(k) \cdot Y(k),$$

and $X \cdot Y = X_\mu Y^\mu = X^\mu Y_\mu$. Also,

$$J_\mu(L; k) \equiv \sum_i -ie_i \int_{L_i} dx_\mu \exp(ikx).$$

The integral along the trajectory L_i is

$$\int_{L_i} dx_\mu \exp(ikx) \equiv \int_{t_1}^{t_2} dt(dx_{i\mu}(t)/dt) \exp(ikx).$$

The $a^*(k)$ and $a(k)$ are the photon creation and annihilation operators:

$$[a(k), a^*(k')] = (2\pi)^3 \delta^3(k - k') 2k_0.$$

The operator $U[L; t_2, t_1]$ acting on the photon vacuum state creates the coherent photon state that is the quantum-theoretic analog of the classical electromagnetic field generated by classical point particles moving on the set of trajectories $L = \{L_i\}$ between times t_1 and t_2.

This $U[L; t_2, t_1]$ can be decomposed into commuting contributions from the various values of k. The general coherent state can be written

$$|q, p> \equiv \exp i(< q \cdot P > - < p \cdot Q >)|0>,$$

where $|0>$ is the photon vacuum state and

$$Q(k) = (a_k + a_k^*)/\sqrt{2}$$

and

$$P(k) = i(a_k - a_k^*)/\sqrt{2}.$$

The $q(k)$ and $p(k)$ are two functions defined (and square integrable) on the mass shell $k^2 = 0$, $k_0 \geq 0$, and the inner product of two coherent states is

$$< q, p|q', p' >= \exp -(< q-q' \cdot q-q' > + < p-p' \cdot p-p' > +2i < p-p' \cdot q+q' >)/4.$$

The coherent states $|q, p>$ can, for various mathematical and physical reasons, be regarded as the "most classical" of the possible states of the electromagnetic quantum field [19]. In order to mock-up the requirements on the possible phenomenal events I assume that each possible phenomenal event corresponds to the actualization of one of these coherent states. In order to represent the idea of a connection of this quantum state to an experiential state I note that each of the modes has a vibratory frequency. Thus each of the coherent states corresponds to a (very high frequency) 'musical chord': each of the notes is being played with a certain strength (it carries a certain energy). One can imagine that the experiencing of this 'chord' constitutes the phenomena. This sort of model has the virtue of being intuitively conceivable, since we do in fact experience musical chords. Also, all experiences can become in an important sense, similiar in kind, and this similarity can extend over all of nature: the experiential aspect of nature can be imagined to be like hearing the 'sound' of the quantum vibrations. The frequency domain thereby becomes pivotal.

In the present model the possibilities for the actualized states are the coherent states $|q, p>$. Thus the ith conscious event is represented by the transition

$$|\Psi_i(t_{i+1}) > \longrightarrow |\Psi_{i+1}(t_{i+1}) >= P_i|\Psi_i(t_{i+1}) >,$$

where $P_i = |q, p; i >< q, p; i|$ is a projection operator.

There is a decomposition of unity

$$I = \prod d^4k(2\pi)^{-4}2\pi\delta^+(k^2) \int dq_k dp_k/\pi$$

$$\times \exp(iq_kP_k - ip_kQ_k)|0_k >< 0_k| \exp -(iq_kP_k - ip_kQ_k).$$

Here meaning can be given by quantizing in a box, so that that the variable k is discretized. Equivalently,

$$I = \int d\mu(q, p)|q, p >< q, p|,$$

where $\mu(q, p)$ is the appropriate measure on the functions q(k) and p(k). Then if the state $|\Psi >< \Psi|$ were to jump to $|q, p >< q, p|$ with probability density $< q, p|\Psi >< \Psi|q, p >$, the resulting mixture would be

$$\int d\mu(q, p)|q, p >< q, p|\Psi >< \Psi|q, p >< q, p|,$$

whose trace is

$$\int d\mu(q, p) < q, p|\Psi >< \Psi|q, p >=< \Psi|\Psi > .$$

Let the state of the electromagnetic field restricted to the modes that represent consciousness be called $|\Psi(t) >$. (Stricly speaking, one must use a density matrix formulation, but the generalization is trivial). Using the decomposition of unity one can write

$$|\Psi(t) >= \int d\mu(q, p)|q, p >< q, p|\Psi(t) > .$$

Hence the state at time t can be represented by the function $< q, p|\Psi(t) >$, which is a complex-valued function over the set of arguments $\{q_1, p_1, q_2, p_2, ..., q_n, p_n\}$, where n is the number of modes associated with $|\Psi >$. This formula expresses the state $|\Psi(t) >$, which in this model is the part of the state of the brain that corresponds to consciousness, as a superposition of states $|q, p >$, each of which is supposed to correspond to a possible phenomenal event.

For each allowed value of k the pair of numbers (q_k, p_k) represents the state of motion of the kth mode of the electromagnetic field. Each of these modes is defined by a particular wave pattern that extends over the whole brain cavity. This pattern is an oscillating structure something like a sine wave or a cosine wave. Each mode is fed by the motions of all of the charged particles in the brain. Thus each mode is a representation of a certain integrated aspect of the activity of the brain.

The state $|q, p >$ represents the conjunction, or collection over the set of all allowed values of k, of the various states $|q_k, p_k >$. The function

$$V(q, p, t) =< q, p|\Psi(t) >< \Psi(t)|q, p >$$

satisfies $0 \leq V(q, p, t) \leq 1$, and it represents, according to orthodox thinking, the probability density that a system that is represented by a general state $|\Psi(t) >$ just before the time t will be observed to be in the classically describable state $|q, p >$ if the observation occurs at time t.

In the absence of interactions, and under certain ideal conditions of confinement, the deterministic normal law of evolution entails that in each mode k there is an independent rotation in the (q_k, p_k) plane with a characteristic angular velocity $\omega_k = k_0$. Due to the effects of the motions of the particles there will be, added to this, a flow of probability that will tend to concentrate the probability in the neighborhoods of a certain set of "optimal" classical-type states $|q, p >$. The reason is that the function of brain dynamics is to produce some single template for action. To be effective this template must be a "classical" state, because, according to orthodox ideas, only 'classical' states can act as dynamically robust control states in the room-temperature brain [20].

According to a semi-classical description of the brain dynamics, only *one* of these classical-type states will be present, but according to quantum theory the state $|\Psi(t) >$ will be a superposition of many such states, unless collapses occurs at lower (i.e., microscopic) levels. The assumption here is that no collapses occur at the lower brain levels.

So in this model the probability will begin to concentrate around various locally optimal coherent states, and hence around the various (generally) isolated points (q,p) in the $2n-$dimensional space at which the quantity

$$V(q,p,t) =< q,p|\Psi_i(t) >< \Psi_i(t)|q,p >$$

reaches a local maximum. Each of these points (q,p) represents a "locally-optimal solution" (at time t) to the search problem.

Eventually there will be an actual event that consists of the experiencing of one of the possible 'chords'. This event will be selected presumably on the basis of some cosmic harmony. But within the confines of contempory science the selection of this state will be governed by 'pure chance', with the probability in the state $|\Psi(t) >$ of the brain for the selected experiencing to be the one corresponding to the state $|q,p >$ being given by the function $V(q,p,t) =< q,p|\Psi(t) >< \Psi(t)|q,p >$.

I shall now use this model to explore the question of the necessity and function of consciousness.

9. The Necessity and Function of Consciousness.

I take it as axiomatic that: 1) human consciousness exists, and 2) human consciousness plays an an essential role in making human bodies behave the way they do.

In the ontological form of quantum theory described in the early sections of this paper our conscious experiencings play a central role: they are the actual things that the propensities specified by the quantum state of the brain are the propensities for. The occurrence of an actual event, which constitutes new knowledge or a new mental conditions, entails a corresponding change in the state of the brain.

However, it would seem logically possible to leave our experiencings out of the causal chain by taking the propensities to be propensities directly for the collapse events themselves, and taking the experiencings to be mere epiphenomemal by-products. But this leads to a technical difficulty that is known as 'the basis problem': what is the set of possible states into which the state of the brain can collapse? In the Copenhagen interpretation these possibilities were defined by our possible experiencings, i.e., by the possible states of our phenomenal knowledge. And in our ontologicalization of the Copenhagen interpretation we have retained this basic role for phenomena: it is the phenomenal possibilities that define the available options, and hence resolve the basis problem.

It might be possible, however, to invent some purely physical criterion that defines the set of possible states into which the state can collapse, and thereby to avoid bringing consciousness into the dynamics. Wigner argued against this possibility by noting that wherever in nature one thing acts on a second there is a reaction back on the first, and it seems unreasonable that consciousness should be an exception. This opinion is represented by our second axiom, which rules out epiphenomenal consciousness.

Regarding the question of the necessity and function of consciousness let us now take as our guide the causal quantum model described above..

In this model the ultimate cause of each selection lies at the cosmic level. In fact, it is an important general truth that the selection process *cannot* be governed by any local process: a local selection process is incompatible with is a the form of the statistical predictions of quantum theory themselves. This is the conclusion of Bell's Theorem, in its strong form [21].

Although contemporary science has as yet little or no understanding of the nature of this nonlocal process, beyond its statistical aspects, we must nevertheless, in the von Neumann/Wigner ontology, take this process into consideration if we are to speak of causes, for in this ontology the nonlocal selection process is the immediate and direct cause of the phenomenal events.

The function of consciousness in our causal quantum mechanical model is that of an intermediary between the physical/bodily process (represented by the Schroedinger-equation-directed evolution of the wave function) and the nonlocal process of selection (represented in our quantum ontology by the chance selections of outcomes of observations). Each possible outcome of an observation is represented by a state $|q, p>$ that is symbolized, for communication with the process of selection, by its 'sound': the hearing of the 'sound' is the experience, and it actualizes the corresponding state.

Generalizing, it would appear that our possible experiencings are the symbolic representations of our possible actions in the common language, or currency, of the universal selection process that quantum theory seems to be telling us is an important player in all physical processes. Our possible experiences are the messages that our brain sends out, and our actual experiences are the ones it receives back. Our experiencings thereby become not epiphenomenal side-shows but rather essential links in the causal chain of events.

Acknowledgements

I thank Pat Hayes and Al Sloman for helpful email dialogs about functionalism.

References.

1. N. Bohr, Atomic Theory and the Description of Nature, Cambridge University Press, Cambridge, 1934. p.1

2. N. Bohr, Atomic Theory and the Description of Nature, Cambridge University Press, Cambridge, 1934. p.1

3. N. Bohr, Atomic Theory and the Description of Nature, Cambridge University Press, Cambridge, 1934. p.18

4. W. Heisenberg, The Representation of Reality in Contemporary Physics, Daedelus, **87**(3), 95-108 (1958)

5. J. von Neumann, Mathematical Foundations of Quantum Mechanics, Princeton Univ. Press, Princeton NJ, (1955), Ch. VI

6. E. Wigner, Remarks on the Mind Body Problem, in The Scientist Speculates, ed. I.J. Good, Heinemann, London, and Basic Books, New York, 1962.

7. Henry P. Stapp, Mind, Matter, and Quantum Mechanics, Springer-Verlag, Heidelberg, Berlin, New York, London, Paris, Tokyo. Ch. 6.

8. R.P. Feynman, The Feynman Lectures in Physics, R.P. Feynman, R.B. Leighton, and M.Sands, Addison-Wesley, (1965) Vol. III, Chapter 21.

9. A. Einstein, in A. Einstein: Philosopher-Scientist, ed. P.A. Schilpp, Tudor, New York, 1951. p.667-673.

10. W. Heisenberg, Physics and Philosophy, Harper Row, New York, 1958, Chapter III.

11. H. Everett III, Rev. Mod. Phys. 29, 463, (1957).

12. W.H. Zurek, in New Techniques and Ideas in Quantum Measurement

Theory, ed, Daniel M. Greenberger, Annals of the New York Academy of Science **480** p.96

13. E. Joos, in New Techniques and Ideas in Quantum Measurement Theory, ed, Daniel M. Greenberger, Annals of the New York Academy of Science **480** p.12

14. M. Gell-Mann and James B. Hartle, Classical Equations for Quantum Systems, UCSBTH-91-15

15. R. Omnes, The Interpretation of Quantum Theory, Princeton University Press, Princeton NJ, 1994, p.348.

16. H.P.Stapp, Phys. Rev. D28, 1386 (1983)

17. T. Kawai and H.P. Stapp, Phys. Rev. D52, 2484-2532, (1995)

18. Philip L. Stocklin and Brian F. Stocklin T.I.T. J. of Life Sci., 1979, Vol 9, pp. 29-51; and Evidence for Endogenous Standing Microwaves as a Substrate for Consciousness. (Paper delivered at the conference "Toward a Scientific Basis for Consciousness, University of Arizona Tucson ,AR, 1994); Physical Basis for Pattern Processing in the Human Brain, 1992

19. R.J. Glauber, in Quantum Optics, S.M. Kay and A. Maitland, eds. Academic Press, London and New York, 1970; T. W. B. Kibble in ibid;
H.P. Stapp, in Quantum Implications: Essays in Honour of David Bohm, B.J. Hiley and F.David Peats eds., Routledge and Paul Kegan Ltd., London and New York, 1987.

20. H.P. Stapp, in Symposium on the Foundations of Modern Physics 1990, P.Lahti and P. Mittelstaedt eds., World Scientific, Singapore. Sec. 3.

21. H.P. Stapp, Strong Versions of Bell's Theorem, Phys. Rev. **49**,3182; N.D. Mermin, Amer. J. Phys. **62**, 880; Ref. 7, p.5.